

# THE POLYNOMIAL CARLESON OPERATOR

VICTOR LIE

*Dedicated to Elias Stein on the  
occasion of his 80<sup>th</sup> birthday celebration.*

**ABSTRACT.** We prove that the generalized Carleson operator  $C_d$  with polynomial phase function is of strong type  $(p, p)$  for  $1 < p < \infty$ , thus answering a question asked by E. Stein ([12],[13]). A key ingredient in this proof is the further extension of the *relational* time-frequency perspective introduced in [9] to the setting of general polynomial phase. Moreover, another important feature of our proof is that, refining the ideas in [4], we provide the first approach to Carleson's Theorem which does not involve exceptional sets. As a consequence, we are able to prove directly, without interpolation techniques, the strong  $L^2$  bound for the Carleson operator, as was anticipated by C. Fefferman (see [4], Remarks).

## 1. Introduction

The main result of this paper is:

**Theorem.** *Let us define the (generalized) polynomial Carleson operator as*

$$(1) \quad C_d f(x) := \sup_{Q \in \mathcal{Q}_d} \left| p.v. \int_{\mathbb{T}} \frac{1}{y} e^{iQ(y)} f(x-y) dy \right| ,$$

*where here  $d \in \mathbb{N}$ ,  $\mathcal{Q}_d$  is the class of all real-coefficient polynomials  $Q$  with  $\deg(Q) \leq d$ , and  $f \in C^1(\mathbb{T})$  with  $\mathbb{T} = [-\frac{1}{2}, \frac{1}{2}]$ . Then for  $1 < p < \infty$  we have*

$$\|C_d f\|_{L^p(\mathbb{T})} \lesssim_{p,d} \|f\|_{L^p(\mathbb{T})} .$$

Thus our Theorem provides the positive answer to:

**Conjecture (E. Stein [12],[13]).** *If  $1 < p < \infty$  then, for any  $d \in \mathbb{N}$ , the polynomial Carleson operator  $C_d$  is of strong type  $(p, p)$ .*

Further on, as one may observe, our Theorem extends Carleson's Theorem on the pointwise convergence of Fourier series, which asserts that  $C_1$  is of

---

*Date:* January 23, 2013.

*2000 Mathematics Subject Classification.* 42A20, 42A50.

*Key words and phrases.* Time-frequency analysis, Carleson's Theorem, polynomial phase.

weak type  $(2, 2)$  ([2], [4], [8]). Also, for  $1 < p < \infty$ ,  $p \neq 2$  we recover (for  $d = 1$ ) the further extension of Hunt ([6]).

On the historical side regarding our result, we mention that if one modifies (1) by taking the supremum only over the class of polynomials (of degree  $d$ ) with no linear term, then one can show the  $L^p$ -boundedness ( $1 < p < \infty$ ) of the resulting operator  $\tilde{C}_d$ . This fact was shown by E. Stein ([12]) for the  $d = 2$  case and for general  $d$  by E. Stein and S. Wainger ([13]). Their method did not require any time-frequency analysis but was based instead on oscillatory-integral techniques. A proof of the boundedness of the operator  $C_2$  was published in [7], but, as it turned out later, this proof was incorrect. Finally, in [9], using a new approach to the time-frequency analysis of the quadratic phase to which we adapted the methods from [4], we proved that  $\|C_2 f\|_1 \lesssim \|f\|_2$ , which together with Stein's maximal principle ([11]) gave us the weak  $L^2$  bound for  $C_2$ .

Passing now to the mathematical aspects of the present paper, we mention here the two main ideas on which our proof is based:

- i) the further extension of the *relational* time-frequency analysis perspective (as described in [9], Section 2) from the quadratic case to the general-polynomial one;
- ii) a new discretization of the family of tiles (Section 5.1.), which provides us the ability of removing the exceptional sets in the decomposition of the Carleson operator, and which, beyond our immediate interest in proving the full range of exponents conjectured by Stein, has some other interesting applications (see Remarks).

Beyond these facts, there will be many other points in our approach (especially Section 7.2.1.) which are inspired from and often follow the intuition and methods developed in [9] for treating the particular case  $d, p = 2$ . These methods were significantly influenced by the powerful geometric and combinatorial ideas presented in [4].

This being said, we briefly elaborate on the two ideas mentioned earlier:

Regarding i), we recall here that one key geometric ingredient in the proof in [9] was to regard the quadratic symmetry from a *relational* perspective. As the name suggests, this perspective stresses the importance of *interactions* between objects rather than simply treating them independently (for further details, see [9]). This approach had as a consequence the splitting of the operator  $C_2$  into “small pieces” with time-frequency portraits (morally) localized near parallelograms (tiles) of area one. In this article, following the above-mentioned perspective, our tiles (that will reflect the time-frequency localization of the “small pieces” of  $C_d$ ) will be some “curved regions” representing neighborhoods<sup>1</sup> of polynomials in the class  $\mathcal{Q}_{d-1}$ .

With respect to ii), the origin of this decomposition lies in our desire to provide a sharp estimate for the  $L^2$ -bound of a family of trees having

---

<sup>1</sup>For the exact meaning of this description, see Section 2.

a uniform, prescribed mass.<sup>2</sup> Such an estimate depends on the counting function associated to the tops of these trees. Now all previously known estimates involved the  $L^\infty$  size of the counting function, and due to this fact one was constrained to remove the sets<sup>3</sup> on which the  $L^\infty$  norm was too large. Based on the intuition provided by the theory of Carleson measures, our belief was that the central role in these estimates should be played not by the  $L^\infty$  but rather by a BMO type norm<sup>4</sup> of the counting function. As a consequence, using a John-Nirenberg argument and an adapted concept of mass, we have organized the family of tiles in a way that allows us to obtain the desired sharp bound.

Next, we should say several words about the structure of our paper:

In Section 2 we present the notations and the general procedure of constructing our tiles, in Section 3 we elaborate on the discretization of our operator  $C_d$ , while Section 4 is dedicated to the study of the interaction between tiles. The key idea in organizing the family of tiles and the Main Proposition are presented in Section 5. Next, in Section 6, we present the main definitions, state Proposition 1 and Proposition 2 and show how to reduce the Main Proposition to the above mentioned results. Section 7 - the most technical one - contains the proofs of the two propositions while Section 8 is dedicated to some final remarks. In the Appendix we include several useful results regarding the distribution and growth of polynomials.

Finally, given that in many respects [4] and [9] can be regarded as a foundation for this paper, when possible, we have chosen to preserve here the notations, definitions and general structure of those earlier works.

**Acknowledgements.** I would like to thank Christoph Thiele for reading the manuscript and giving useful feedback.

## 2. Notations and construction of the tiles

As mentioned in the introduction, we denote by  $\mathcal{Q}_d$  the class of all real polynomials of degree at most  $d$ . If not specified,  $q$  will always designate an element of  $\mathcal{Q}_{d-1}$ , while  $Q$  will refer to an element of  $\mathcal{Q}_d$ . When appearing together in a proof  $q$  will designate the derivative of  $Q$ .

Take now the canonical dyadic grids in  $[0, 1] = \mathbb{T}$  and in  $\mathbb{R}$ .<sup>5</sup> Throughout the paper the letters  $I$  and  $J$  will refer to dyadic intervals corresponding to the grid in  $\mathbb{T}$  while  $\alpha^1, \dots, \alpha^d$  will represent dyadic intervals associated with the grid in  $\mathbb{R}$ . Now, if  $I$  is any (dyadic) interval we denote by  $c(I)$  the center of  $I$ . Let  $I_r$  be the “right brother” of  $I$ , with  $c(I_r) = c(I) + |I|$  and  $|I_r| = |I|$ ;

---

<sup>2</sup>See Section 5 for definitions.

<sup>3</sup>These “exceptional” sets caused a series of technical problems in all the earlier works regarding the  $L^p$ -boundedness of the Carleson operator, and they were responsible for the lack of a direct approach for providing strong  $L^2$  bounds.

<sup>4</sup>See Section 5.1. for the precise definition.

<sup>5</sup>The reader should not be confused by the fact that, depending on our convenience, the symbol  $\mathbb{T}$  may refer to a different unit interval from that mentioned in the statement of our Theorem.

similarly, the “left brother” of  $I$  will be denoted  $I_l$  with  $c(I_l) = c(I) - |I|$  and  $|I_l| = |I|$ . If  $a > 0$  is some real number, by  $aI$  we mean the interval with the same center  $c(I)$  and with length  $|aI| = a|I|$ ; the same conventions apply to intervals  $\{\alpha^k\}_k$ .

A *tile*  $P$  is a  $(d+1)$ -tuple of dyadic (half open) intervals, *i.e.*  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$ , such that  $|\alpha^j| = |I|^{-1}$ ,  $j \in \{1, \dots, d\}$ . The collection of all tiles  $P$  will be denoted by  $\mathbb{P}$ .

Now, for each tile  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$  we will associate a geometric (time-frequency) representation, denoted  $\hat{P}$ . The exact procedure is as follows: for  $I$  as before we first set  $x_I = (x_I^1, x_I^2, \dots, x_I^d) \in \mathbb{T}^d$  to be the  $d$ -tuple defined as follows:  $x_I^1, x_I^2$  are the endpoints of the interval  $I$ ,  $x_I^3 = \frac{x_I^1 + x_I^2}{2}$ ,  $x_I^4 = \frac{x_I^1 + x_I^3}{2}$ , then  $x_I^5 = \frac{x_I^3 + x_I^2}{2}$ , and inductively (in the obvious manner) we continue this procedure until we reach the  $d$ -th coordinate. Then, define

$$\mathcal{Q}_{d-1}(P) = \{q \in \mathcal{Q}_{d-1} \mid q(x_I^j) \in \alpha^j \ \forall j \in \{1, \dots, d\}\}.$$

We will say that  $q \in P$  iff  $q \in \mathcal{Q}_{d-1}(P)$ . Finally, we set

$$(2) \quad \hat{P} = \{(x, q(x)) \mid x \in I \ \& \ q \in P\}.$$

The collection of all geometric tiles  $\hat{P}$  will be denoted with  $\hat{\mathbb{P}}$ .

In the following we will also work with dilates of our tiles: for  $a > 0$  and  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$  we set  $aP := [a\alpha^1, a\alpha^2, \dots, a\alpha^d, I]$ . Similarly, we write

$$a\hat{P} := \widehat{aP} = \{(x, q(x)) \mid x \in I \ \& \ q \in \mathcal{Q}_{d-1}(aP)\}.$$

Also, if  $\mathcal{P} \subseteq \mathbb{P}$  then by convention  $a\mathcal{P} := \{aP \mid P \in \mathcal{P}\}$ ; similarly, if  $\hat{\mathcal{P}} \subseteq \hat{\mathbb{P}}$  then  $\widehat{a\mathcal{P}} := \{a\hat{P} \mid \hat{P} \in \hat{\mathcal{P}}\}$ .

For each tile  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P}$  we associate the “central polynomial”  $q_P \in \mathcal{Q}_{d-1}$  given by the Lagrange interpolation polynomial:

$$(3) \quad q_P(y) := \sum_{j=1}^d \frac{\prod_{\substack{k=1 \\ k \neq j}}^d (y - x_I^k)}{\prod_{\substack{k=1 \\ k \neq j}}^d (x_I^j - x_I^k)} c(\alpha^j).$$

For  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$  we denote the collection of its neighbors by  $N(P) = \{P' = [\alpha^{1'}, \alpha^{2'}, \dots, \alpha^{d'}, I] \mid \alpha^{k'} \in \{\alpha^k, \alpha_r^k, \alpha_l^k\} \ \forall k \in \{1, \dots, d\}\}$ .

For any dyadic interval  $I \subseteq [0, 1]$ , define the (non-dyadic) intervals

$$I_r^* = [c(I) + \frac{7}{2}|I|, c(I) + \frac{11}{2}|I|) \quad \& \quad I_l^* = [c(I) - \frac{11}{2}|I|, c(I) - \frac{7}{2}|I|)$$

and set  $I^* = I_r^* \cup I_l^*$  and  $\tilde{I} = 13I$ . Similarly, for  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$ , we define

$$\widehat{P}_r^* = \{(x, q(x)) \mid x \in I_r^* \ \& \ q \in P\}$$

and further repeat the same procedure for  $\widehat{P}_l^*$ ,  $\widehat{P}^*$  and  $\widehat{\tilde{P}}$ .

Throughout the paper  $p$  will be the index of the Lebesgue space  $L^p$  and, unless otherwise mentioned, will obey  $1 < p < \infty$ . Also,  $p'$  will be its Hölder conjugate (*i.e.*  $\frac{1}{p} + \frac{1}{p'} = 1$ ), while  $p^* =_{\text{def}} \min(p, p')$ .

For  $f \in L^p(\mathbb{T})$ , we denote by

$$Mf(x) = \sup_{x \in I} \frac{1}{|I|} \int_I |f|$$

the Hardy-Littlewood maximal function associated to  $f$ .

If  $\{I_j\}$  is a collection of pairwise disjoint intervals in  $[0, 1]$  and  $\{E_j\}$  a collection of sets such that for a fixed  $\delta \in (0, 1)$

$$(4) \quad E_j \subset I_j \quad \& \quad \frac{|E_j|}{|I_j|} \leq \delta \quad \forall j \in \mathbb{N},$$

then we denote

$$(5) \quad M_\delta f(x) := \begin{cases} \sup_{I \supset I_j} \frac{1}{|I|} \int_I |f|, & \text{if } x \in E_j \\ 0, & \text{if } x \notin E_j \end{cases}.$$

For  $A, B > 0$  we say  $A \lesssim B$  (resp.  $A \gtrsim B$ ) if there exists an absolute constant  $C > 0$  such that  $A < CB$  (resp.  $A > CB$ ); if the constant  $C$  depends on some quantity  $\delta > 0$  then we may write  $A \lesssim_\delta B$ . If  $C^{-1}A < B < CA$  for some (positive) absolute constant  $C$  then we write  $A \approx B$ .

As in [9], for  $x \in \mathbb{R}$  we set  $\lceil x \rceil := \frac{1}{1+|x|}$ .

The parameters  $\eta = \eta(d)$ ,  $c(d)$  (designating positive numbers depending on  $d$ ) and  $c$  (standing for a large positive number) are allowed to change from line to line.

In what follows, for notational simplicity, we will refer to the operator  $C_d$  as  $T$ .

### 3. Discretization

We first express  $T$  in terms of its elementary building blocks:

$$Tf(x) = \sup_{a_1, \dots, a_d \in \mathbb{R}} |M_{1,a_1} \dots M_{d,a_d} H M_{1,a_1}^* \dots M_{d,a_d}^* f(x)| = \sup_{Q \in \mathcal{Q}_d} |T_Q f(x)|,$$

where  $\{M_{j,a_j}\}_{j \in \{1, \dots, d\}}$  is the family of (generalized) modulations given by

$$M_{j,a_j} f(x) := e^{ia_j x^j} f(x) \quad j \in \{1, \dots, d\}$$

(here  $f \in L^p$ ,  $a_j \in \mathbb{R}$  &  $x \in \mathbb{T}$ ) and

$$T_Q f(x) = \int_{\mathbb{T}} \frac{1}{y} e^{i(Q(x) - Q(x-y))} f(x-y) dy$$

with  $Q \in \mathcal{Q}_d$  given by  $Q(y) = \sum_{j=1}^d a_j y^j$ .

Equivalently

$$T_Q f(x) = \int_{\mathbb{T}} \frac{1}{x-y} e^{i(\int_y^x q)} f(y) dy,$$

where here, as mentioned in the previous section,  $q$  stands for the derivative of  $Q$ .

Now linearizing  $T$  we write

$$Tf(x) = T_{Q_x}f(x) = \int_{\mathbb{T}} \frac{1}{x-y} e^{i(\int_y^x q_x)} f(y) dy ,$$

where  $Q_x(y) := \sum_{j=1}^d a_j(x)y^j$  with  $\{a_j(\cdot)\}_{j \in \{1, \dots, d\}}$  measurable functions and  $q_x$  is the derivative of  $Q_x$  (i.e.  $q_x(t) = \frac{d}{dt}Q_x(t)$ ).

Further, proceeding as in [4] and [9], we define  $\psi$  to be an odd  $C^\infty$  function such that  $\text{supp } \psi \subseteq \{y \in \mathbb{R} \mid 2 < |y| < 8\}$  and

$$\frac{1}{y} = \sum_{k \geq 0} \psi_k(y) \quad \forall \ 0 < |y| < 1 ,$$

where by definition  $\psi_k(y) := 2^k \psi(2^k y)$  (with  $k \in \mathbb{N}$ ). Using this, we deduce that

$$Tf(x) = \sum_{k \geq 0} T_k f(x) := \sum_{k \geq 0} \int_{\mathbb{T}} e^{i(\int_y^x q_x)} \psi_k(x-y) f(y) dy .$$

Now for each  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P}$  let  $E(P) = \{x \in I \mid q_x \in P\}$ . Also, if  $|I| = 2^{-k}$  ( $k \geq 0$ ), we define the operators  $T_P$  on  $L^2(\mathbb{T})$  by

$$T_P f(x) = \left\{ \int_{\mathbb{T}} e^{i(\int_y^x q_x)} \psi_k(x-y) f(y) dy \right\} \chi_{E(P)}(x) .$$

As expected, if  $\mathbb{P}_k := \{P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P} \mid |I| = 2^{-k}\}$ , for fixed  $k$  the  $E(P)$  form a partition of  $[0, 1]$ , and so

$$T_k f(x) = \sum_{P \in \mathbb{P}_k} T_P f(x) .$$

Consequently, we have

$$Tf(x) = \sum_{k \geq 0} T_k f(x) = \sum_{P \in \mathbb{P}} T_P f(x) .$$

This ends our decomposition.

Finally, note that (as in [9]) we may assume that

$$\text{supp } \psi \subseteq \{y \in \mathbb{R} \mid 4 < |y| < 5\} .$$

Consequently, for a tile  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$ , the associated operator has the properties

$$\text{supp } T_P \subseteq I \quad \& \quad \text{supp } T_P^* \subseteq \{x \mid 3|I| \leq \text{dist}(x, I) \leq 5|I|\} = I^* ,$$

where here  $T_P^*$  denotes the adjoint of  $T_P$ .

Also, in what follows, (splitting  $\mathbb{P} = \bigcup_{j=0}^{D-1} \bigcup_{k \geq 0} \mathbb{P}_{kD+j}$  where  $D$  is the smallest integer larger than  $2d \log_2(2d)$ ) we can suppose that if  $P_j = [\alpha_j^1, \alpha_j^2, \dots, \alpha_j^d, I_j] \in \mathbb{P}$  with  $j \in \{1, 2\}$  such that  $|I_1| \neq |I_2|$ , then  $|I_1| \leq 2^{-D} |I_2|$  or  $|I_2| \leq 2^{-D} |I_1|$ .

#### 4. Quantifying the interactions between tiles

In this section we will focus on the behavior of the expression

$$(6) \quad \left| \langle T_{P_1}^* f, T_{P_2}^* g \rangle \right| .$$

Before this, we will need to introduce some quantitative concepts that are adapted to the information offered by the localization of  $\{T_{P_j}\}$ .

##### 4.1. Properties of $T_P$ and $T_P^*$

For  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P}$  with  $|I| = 2^{-k}$ ,  $k \in \mathbb{N}$ , we have

$$(7) \quad \begin{aligned} T_P f(x) &= \left\{ \int_{\mathbb{T}} e^{i(\int_y^x q_x)} \psi_k(x-y) f(y) dy \right\} \chi_{E(P)}(x) , \\ T_P^* f(x) &= \int_{\mathbb{T}} e^{-i(\int_x^y q_y)} \psi_k(y-x) (\chi_{E(P)} f)(y) dy . \end{aligned}$$

Based on the *relational* approach developed in [9] we have:

$$(8) \quad \begin{aligned} &- \text{the time-frequency localization of } T_P \text{ is “morally” given by the tile } \hat{P}; \\ &- \text{the time-frequency localization of } T_P^* \text{ is “morally” given by the (bi)tile } \widehat{P}^*. \end{aligned}$$

(Remark that, due to Lemma C (see the Appendix), one may think of  $\hat{P}$  as the roughly  $|I|^{-1}$  neighborhood of the graph of the “central polynomial”  $q_P$  restricted to the interval  $I$ .)

##### 4.2. Factors of a tile

For a tile  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I]$  we define two quantities:

a) an *absolute* one (which may be regarded as a self-interaction); we define the **density (analytic) factor of  $P$**  to be the expression

$$(9) \quad A_0(P) := \frac{|E(P)|}{|I|} .$$

Notice that  $A_0(P)$  determines the  $L^2$  operator norm of  $T_P$ .

b) a *relative* one (interaction of  $P$  or  $\hat{P}$  with an exterior object) which is of a geometric nature.

Suppose first that we are given  $q \in \mathcal{Q}_{d-1}$  and  $J$  a dyadic interval; we introduce the quantity

$$\Delta_q(J) := \frac{\text{dist}^J(q, 0)}{|J|^{-1}} ,$$

where we use the notations (for  $q_1, q_2 \in \mathcal{Q}_{d-1}$ )

$$\text{dist}^A(q_1, q_2) = \sup_{y \in A} \{\text{dist}_y(q_1, q_2)\} \quad \& \quad \text{dist}_y(q_1, q_2) = |q_1(y) - q_2(y)| .$$

Now we define the **geometric factor of  $P$  ( $\hat{P}$ ) with respect to  $q$**  to be the term

$$[\Delta_q(P)] \left( = \frac{1}{1 + |\Delta_q(P)|} \right) ,$$

where

$$(10) \quad \Delta_q(P) := \inf_{q_1 \in P} \Delta_{q-q_1}(I_P) .$$

### 4.3. The resulting estimates

We conclude this section by observing how the above quantities relate in controlling the interaction in (6).

As expected, we need to quantify the relative position of  $\widehat{P}_1^*$  with respect to  $\widehat{P}_2^*$ . (We consider only the nontrivial case  $I_{P_1}^* \cap I_{P_2}^* \neq \emptyset$ ; also, throughout this section we suppose that  $|I_1| \geq |I_2|$ .)

**Definition 1.** *Given two tiles  $P_1$  and  $P_2$ , we define the **geometric factor of the pair  $(P_1, P_2)$**  by*

$$[\Delta(P_1, P_2)] ,$$

where

$$\Delta(P_1, P_2) = \Delta_{1,2} := \frac{\sup_{y \in I_2} \left\{ \inf_{\substack{q_1 \in P_1 \\ q_2 \in P_2}} \text{dist}_y(q_1, q_2) \right\}}{|I_2|^{-1}} .$$

With these notations, remark (using the results in the Appendix) that we have

$$[\Delta_{1,2}] \approx_d \max \left\{ \left[ \Delta_{q_{P_1}}(P_2) \right], \left[ \Delta_{q_{P_2}}(P_1) \right] \right\} .$$

For  $P_1$  and  $P_2$  as above, we define the “interaction polynomial”

$$q_{1,2} := q_{P_1} - q_{P_2} .$$

Fix now an interval (not necessarily dyadic)  $\bar{J} \subseteq \mathbb{T}$ , a polynomial  $q \in \mathcal{Q}_{d-1}$  and three positive constants  $\eta, v, w$ . In what follows we will present a general procedure for constructing two types of critical sets associated with  $\bar{J}$ ,  $q$ ,  $\eta$ ,  $v$  and  $w$ , denoted  $\mathcal{I}_s(\eta, v, q, \bar{J})$  and  $\mathcal{I}_c(\eta, w, q, \bar{J})$ .

Suppose for the moment that  $q \notin \mathcal{Q}_0$ ; let  $J$  be the largest<sup>6</sup> dyadic interval contained in  $\bar{J}$ . We define

$$\mathcal{M}_q^\eta(\bar{J}) = \{x \in \bar{J} \mid x \text{ is a local minimum for } |q| \text{ \& } |q|(x) < \eta\} .$$

Now since  $q \in \mathcal{Q}_{d-1} \setminus \mathcal{Q}_0$  we have that  $\mathcal{M}_q^\eta(\bar{J})$  is a finite set of the form  $\mathcal{M}_q^\eta(\bar{J}) = \{x_j\}_{j \in \{1, \dots, r\}}$  with  $r \leq 2d$ . Without loss of generality we suppose that the  $x_j$  are arranged in increasing order.

---

<sup>6</sup>If there are two such intervals just pick either of them.



Further, for each  $j \in \{1, \dots, r\}$  define  $\tilde{q}_j(x) := q(x) - q(x_j)$  and construct the dyadic intervals  $I_j^1$ ,  $I_j^2$  and  $I_j^3$  as follows:  $I_j^1$  is the smallest dyadic interval  $I$  for which  $x_j \in I$  and  $\Delta_{\tilde{q}_j}(I) > c(d)v$ ;  $I_j^2$  is the smallest dyadic interval  $I$  whose left endpoint is equal to the right endpoint of  $I_j^1$  and for which  $\Delta_{\tilde{q}_j}(I) > c(d)v$ ; similarly,  $I_j^3$  is the smallest dyadic interval  $I$  whose right endpoint is equal to the left endpoint of  $I_j^1$  and for which  $\Delta_{\tilde{q}_j}(I) > c(d)v$ .

Now set

$$\mathcal{S}_j(\eta, v, q, \bar{J}) = \bigcup_{k=1}^3 I_j^k$$

and define

$$\mathcal{S}(\eta, v, q, \bar{J}) = \bigcup_{j=1}^r \mathcal{S}_j(\eta, v, q, \bar{J}).$$

Also set

$$\mathcal{C}_j(\eta, w, q, \bar{J}) = [x_j - w, x_j + w] \cap \bar{J}$$

and further take

$$\mathcal{C}(\eta, w, q, \bar{J}) = \bigcup_{j=1}^r \mathcal{C}_j(\eta, w, q, \bar{J}).$$

We now need to do one more step before ending our construction; suppose that  $A \subseteq \bar{J}$  is a finite union of (closed) intervals:  $A = \bigcup_{j=1}^l A_j$  with  $l \in \mathbb{N}$ ,  $A_j = [u_j, v_j]$  (pairwise disjoint) and  $\{u_j\}$  monotone increasing. Then, setting  $A_0 = A_{l+1} = \emptyset$  we define

$$\mathcal{E}(\bar{J}, A) = \left( \bigcup_{j=1}^l A_j \right) \cup \left( \bigcup_{\substack{j \in \{1, \dots, l+1\} \\ |C_j| < |A_{j-1}| \text{ or } |C_j| < |A_j|}} C_j \right),$$

where here the intervals  $C_j$  obey the partition condition

$$\bar{J} = \bigcup_{j=1}^{l+1} (A_j \cup C_j).$$

Finally, if  $q \notin \mathcal{Q}_0$ , define

$$\mathcal{I}_s(\eta, v, q, \bar{J}) = \mathcal{E}(\bar{J}, \mathcal{S}(\eta, v, q, \bar{J}))$$

and

$$\mathcal{I}_c(\eta, w, q, \bar{J}) = \mathcal{E}(\bar{J}, \mathcal{C}(\eta, w, q, \bar{J})).$$

Otherwise, if  $q \in \mathcal{Q}_0$ , just set  $\mathcal{I}_s(\eta, v, q, \bar{J}) = \mathcal{I}_c(\eta, w, q, \bar{J}) = \emptyset$ .

Fix  $\epsilon_0 \in (0, 1)$ . Set  $w(\bar{J}) = c(d)|J|[\Delta_q(J)]^{\frac{1}{d}-\epsilon_0}$ ,  $v(\bar{J}) = c(d)[\Delta_q(J)]^{-2\epsilon_0}$  and  $\eta(\bar{J}) = c(d)v(\bar{J})w(\bar{J})^{-1}$ .

Then, using the results in the Appendix, we deduce

$$(11) \quad \mathcal{I}_s(\eta(\bar{J}), v(\bar{J}), q, \bar{J}) \subseteq \mathcal{I}_c(\eta(\bar{J}), w(\bar{J}), q, \bar{J}).$$

We now define the  $(\epsilon_0)$ -**critical intersection set**  $I_{1,2}$  of the pair  $(P_1, P_2)$  as

$$(12) \quad I_{1,2} := \mathcal{I}_c(\eta_{1,2}, w_{1,2}, q_{1,2}, \tilde{I}_1 \cap \tilde{I}_2),$$

where  $\eta_{1,2} := \eta(\tilde{I}_1 \cap \tilde{I}_2)$  and  $w_{1,2} := w(\tilde{I}_1 \cap \tilde{I}_2)$ .

Notice that, based on (11) and Lemma C of the Appendix, we have that

$$(13) \quad \bigcup_{\substack{q_j \in P_j \\ j \in \{1,2\}}} \left\{ y \in \tilde{I}_2 \mid \frac{|q_1(y) - q_2(y)|}{|I_2|^{-1}} \leq \lceil \Delta_{1,2} \rceil^{-\frac{1}{d} - \epsilon_0} \right\} \subseteq I_{1,2}.$$

Now using (13) together with the principle of (non-)stationary phase, one deduces the following:

**Lemma 0.** *Let  $P_1, P_2 \in \mathbb{P}$ ; then we have*

$$(14) \quad \left| \int \tilde{\chi}_{I_{1,2}^c} T_{P_1}^* f \overline{T_{P_2}^* g} \right| \lesssim_{n,d,\epsilon_0} \lceil \Delta(P_1, P_2) \rceil^n \frac{\int_{E(P_1)} |f| \int_{E(P_2)} |g|}{\max(|I_1|, |I_2|)} \quad \forall n \in \mathbb{N},$$

$$(15) \quad \left| \int_{I_{1,2}} T_{P_1}^* f \overline{T_{P_2}^* g} \right| \lesssim_d \lceil \Delta(P_1, P_2) \rceil^{\frac{1}{d} - \epsilon_0} \frac{\int_{E(P_1)} |f| \int_{E(P_2)} |g|}{\max(|I_1|, |I_2|)},$$

where  $\tilde{\chi}_{I_{1,2}^c}$  is a smooth variant of the corresponding cut-off.

Applying the same methods for the limiting case  $\epsilon_0 = 0$ , we obtain

$$(16) \quad \|T_{P_1} T_{P_2}^*\|_2^2 \lesssim_d \min \left\{ \frac{|I_2|}{|I_1|}, \frac{|I_1|}{|I_2|} \right\} \lceil \Delta(P_1, P_2) \rceil^{\frac{2}{d}} A_0(P_1) A_0(P_2).$$

*Proof.* We first notice that relation (15) is straightforward; indeed, to see this we just use the relation

$$|T_{P_j}^* f| \lesssim \frac{\int_{E(P_j)} |f|}{|I_j|} \chi_{I_j^*} \quad \forall j \in \{1, 2\}$$

together with the definition of  $I_{1,2}$ .

We now turn our attention towards (14). First, for notational convenience we set  $\varphi = \tilde{\chi}_{I_{1,2}^c}$ ; with this, we have:

$$\begin{aligned} \int \varphi T_{P_1}^* f \overline{T_{P_2}^* g} &= \int f \overline{T_{P_1}(\varphi T_{P_2}^* g)} \\ &= \int \int (f \chi_{E(P_1)})(x) (\bar{g} \chi_{E(P_2)})(s) \mathcal{K}(x, s) dx ds, \end{aligned}$$

where

$$\mathcal{K}(x, s) = \int e^{i[\int_y^s q_s - \int_y^x q_x]} \psi_{k_1}(x - y) \varphi(y) \psi_{k_2}(y - s) dy.$$

(Here we use the conventions  $|I_1| = 2^{-k_1}$ ,  $|I_2| = 2^{-k_2}$  with  $k_2 \geq k_1$  positive integers.)

Now making the change of variables  $y = |I_2| u$  and using the definitions of  $I_{1,2}$  and  $\varphi$ , we deduce that

$$|\mathcal{K}(x, s)| \lesssim |I_1|^{-1} \left| \int_{\mathbb{T}} e^{i\phi(u)} r(u) du \right|$$

with  $r \in C_0^\infty(\mathbb{R})$  such that  $|\partial^l r(u)| \lesssim_d ([\Delta(P_1, P_2)]^{\epsilon_0 - \frac{1}{d}})^l$  ( $l \in \mathbb{N}$ ) and  $\|\partial\phi\|_{L^\infty(\text{supp } r)} \gtrsim_d [\Delta(P_1, P_2)]^{-\epsilon_0 - \frac{1}{d}}$ . Using the non-stationary phase principle we thus obtain (14).

For (16), we set  $\epsilon_0 = 0$  in the previous argument. □

## 5. The proof of the main theorem

### 5.1. A key ingredient - organizing the family of tiles.

In this section we will recursively partition the set of all tiles  $\mathbb{P}$  into families of tiles with some special properties. More precisely, using induction, we will show that

$$(17) \quad \mathbb{P} = \bigcup_n \mathbb{P}_n,$$

such that, roughly speaking, for each family  $\mathbb{P}_n$  we have that

- the tiles inside have a uniform density factor;
- the counting function(s) associated with our family is (are) under “good” control.

To make this precise, we need to introduce the following

**Definition 2.** Let  $\mathcal{A}$  be a (finite) union of dyadic intervals in  $[0, 1]$  and  $\mathcal{P}$  be a finite family of tiles. For  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}$  with  $I \subseteq \mathcal{A}$  we define the **mass** of  $P$  relative to the set of tiles  $\mathcal{P}$  and the set  $\mathcal{A}$  as being

$$(18) \quad A_{\mathcal{P}, \mathcal{A}}(P) := \sup_{\substack{P' = [\alpha^{1'}, \alpha^{2'}, \dots, \alpha^{d'}, I'] \in \mathcal{P} \\ I \subseteq I' \subseteq \mathcal{A}}} \frac{|E(P')|}{|I'|} [\Delta(2P, 2P')]^N$$

where  $N$  is a fixed large natural number.

Let us firstly see the procedure for constructing the family  $\mathbb{P}_1$ .

For this, let  $\mathcal{P}_1^{0, \max}$  be the collection of maximal tiles  $P \in \mathbb{P}$  with  $\frac{|E(P)|}{|I|} \geq \frac{1}{2}$ . Also define the initial set of the time intervals of these maximal tiles as  $\mathcal{I}_1^0 := \{I \mid P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}_1^{0, \max}\}$ .

Notice that since  $\mathcal{P}_1^{0, \max}$  is formed by disjoint tiles we have that  $\mathcal{C}_1^0 := \sum_{P \in \mathcal{P}_1^{0, \max}} \chi_{E(P)}$  obeys  $\mathcal{C}_1^0(x) \leq 1$ . This further implies that the counting

function  $\mathcal{N}_1^0 := \sum_{I \in \mathcal{I}_1^0} \chi_I$  verifies the relation

$$(19) \quad \|\mathcal{N}_1^0\|_{BMO_C} := \sup_{\substack{J \text{ dyadic} \\ J \subseteq [0,1]}} \frac{\sum_{\substack{I \subseteq J \\ I \in \mathcal{I}_1^0}} |I|}{|J|} \leq 2.$$

Thus  $\mathcal{N}_1^0 \in BMO_D(\mathbb{R})$ .

Applying now the the John-Nirenberg inequality, we have

$$(20) \quad |\{x \in J \mid |\mathcal{N}_1(x) - \frac{\int_J \mathcal{N}_1^0}{|J|}| > \gamma\}| \lesssim |J| e^{-c \frac{\gamma}{\|\mathcal{N}_1^0\|_{BMO_D(\mathbb{R})}}}.$$

Next, we notice that  $\|\mathcal{N}_1^0\|_{BMO_D(\mathbb{R})} \leq 2\|\mathcal{N}_1^0\|_{BMO_C}$ . This further implies that for  $\gamma > c\|\mathcal{N}_1^0\|_{BMO_C}$  we have that

$$(21) \quad |\{x \in J \mid \sum_{\substack{I \subseteq J \\ I \in \mathcal{I}_1^0}} \chi_I(x) > \gamma\}| \lesssim |J| e^{-c}.$$

Let us define

$$A_1^1 := \{x \in [0, 1] \mid \sum_{\substack{I \subseteq [0,1] \\ I \in \mathcal{I}_1^0}} \chi_I(x) > c\|\mathcal{N}_1^0\|_{BMO_C}\}.$$

Applying now (21) we deduce that  $|A_1^1| \leq e^{-c}$ .

Further let  $\mathcal{P}_1^{1,max}$  be the collection of maximal tiles  $P \in \mathbb{P}$  with  $\frac{|E(P)|}{|I_P|} \geq \frac{1}{2}$  and  $I_P \subseteq A_1^1$ . Now as before, we define the collection of time-intervals of maximal tiles as  $\mathcal{I}_1^1 := \{I \mid P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}_1^{1,max}\}$ .

For  $\mathcal{C}_1^1 := \sum_{P \in \mathcal{P}_1^{1,max}} \chi_{E(P)}$  we again have  $\mathcal{C}_1^1(x) \leq 1$  which implies that the counting function  $\mathcal{N}_1^1 := \sum_{I \in \mathcal{I}_1^1} \chi_I$  is in  $BMO_D(\mathbb{R})$  and moreover that  $\|\mathcal{N}_1^1\|_{BMO_C} \leq 2$ .

Using as before John-Nirenberg inequality, for  $\gamma > c\|\mathcal{N}_1^1\|_{BMO_C}$ , we have

$$(22) \quad |\{x \in J \mid \sum_{\substack{I \subseteq J \\ I \in \mathcal{I}_1^1}} \chi_I(x) > \gamma\}| \lesssim |J| e^{-c}.$$

Thus, defining now the set

$$A_1^2 := \{x \in [0, 1] \mid \sum_{\substack{I \subseteq [0,1] \\ I \in \mathcal{I}_1^1}} \chi_I(x) > c\|\mathcal{N}_1^1\|_{BMO_C}\},$$

we have that  $|A_1^2| \leq e^{-c}|A_1^1|$ .

Continuing by induction, at the step  $k$  we will be able to construct the analogue sets  $A_1^k$ ,  $\mathcal{P}_1^{k,max}$ ,  $\mathcal{I}_1^k$  and respectively the counting function  $\mathcal{N}_1^k$ . This process will end in a finite number of steps since the family  $\mathbb{P}$  is finite.

Let us now define the 1-maximal set of tiles  $\mathcal{P}_1^{max} := \bigcup_k \mathcal{P}_1^{k,max}$ , the collection of the time-intervals  $\mathcal{I}_1 := \bigcup_k \mathcal{I}_1^k$  and finally the counting function  $\mathcal{N}_1 := \sum_{I \in \mathcal{I}_1} \chi_I$ .

Now as a consequence of the above construction we deduce that:

$$(23) \quad \|\mathcal{N}_1\|_{BMO_C} \lesssim \max_k \|\mathcal{N}_1^k\|_{BMO_C}.$$

Moreover, for any  $l < k$ , we have that  $A_1^k \subset A_1^l$  with

$$(24) \quad |A_1^k| \leq e^{-(k-l)c} |A_1^l|.$$

Next, define

$$\mathcal{P}_1^0 := \{P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P} \mid I \not\subseteq A_1^1 \text{ \& } A_{\mathbb{P}, [0,1]}(P) \in [2^{-1}, 2^0)\}$$

and further, by induction, construct

$$\mathcal{P}_1^k := \left\{ P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P} \mid \begin{array}{l} I \not\subseteq A_1^{k+1}, I \subseteq A_1^k \\ A_{\mathbb{P}, A_1^k}(P) \in [2^{-1}, 2^0) \end{array} \right\}.$$

Finally set

$$\mathbb{P}_1 := \bigcup_k \mathcal{P}_1^k.$$

Here the construction of the 1-mass set ends.

With this done, suppose now that we have constructed the sets  $\{\mathbb{P}_k\}_{k < n}$  and let us see how we define the set  $\mathbb{P}_n$ .

First step consists from selecting the family  $\mathcal{P}_n^{0,max}$  of the maximal tiles  $P \in \mathbb{P} \setminus \bigcup_{k < n} \mathbb{P}_k$  with  $\frac{|E(P)|}{|I_P|} \geq 2^{-n}$ . After that, we collect the time-intervals of these maximal tiles into the set  $\mathcal{I}_n^0$  and form with them the counting function

$$\mathcal{N}_n^0 := \sum_{I \in \mathcal{I}_n^0} \chi_I.$$

Next, by using John-Nirenberg inequality we remark that the set

$$A_n^1 := \{x \in [0, 1] \mid \sum_{I \in \mathcal{I}_n^0} \chi_I(x) > cn \|\mathcal{N}_n^0\|_{BMO_C}\},$$

has the measure  $|A_n^1| \leq e^{-100n}$ .

Further, we construct  $\mathcal{P}_n^{1,max}$  to be the collection of maximal tiles  $P \in \mathbb{P} \setminus \bigcup_{k < n} \mathbb{P}_k$  with  $\frac{|E(P)|}{|I_P|} \geq 2^{-n}$  and  $I_P \subseteq A_n^1$ . Also, as before, define  $\mathcal{I}_n^1 := \{I \mid P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}_n^{1,max}\}$ , the counting function  $\mathcal{N}_n^1 := \sum_{I \in \mathcal{I}_n^1} \chi_I$  and the exceptional set

$$A_n^2 := \{x \in [0, 1] \mid \sum_{I \in \mathcal{I}_n^1} \chi_I(x) > cn \|\mathcal{N}_n^1\|_{BMO_C}\}.$$

Proceeding by induction, at the end of the day, we will have constructed the collection of sets of maximal tiles  $\{\mathcal{P}_n^{k,max}\}_k$ , the collection of sets representing the time-intervals -  $\{\mathcal{I}_n^k\}_k$ , the collection of counting functions  $\{\mathcal{N}_n^k\}_k$  and finally the level sets  $\{A_n^k\}_k$ .

Reached in this point we state the following important consequences of our construction:

$$(25) \quad |A_n^k| \leq e^{-100|k-l|n} |A_n^l| ,$$

$$(26) \quad \sup_k \|\mathcal{N}_n^k\|_{BMO_C} \leq 2^n \text{ and } \sup_k \|\mathcal{N}_n^k\|_{L^\infty(A_n^k \setminus A_n^{k+1})} \lesssim n 2^n .$$

Moreover, if we set the counting function

$$\mathcal{N}_n := \sum_k \mathcal{N}_n^k ,$$

we also have

$$(27) \quad \|\mathcal{N}_n\|_{BMO_C} \lesssim 2^n .$$

With this done, for each  $k \in \mathbb{N}$ , we define

$$(28) \quad \mathcal{P}_n^k := \left\{ P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \mid \begin{array}{l} I \subseteq A_n^k, I \not\subseteq A_n^{k+1} \text{ and} \\ A_{\mathbb{P} \setminus \bigcup_{j < n} \mathbb{P}_j, A_n^k}(P) \in [2^{-n}, 2^{-n+1}) \end{array} \right\} ,$$

and set

$$(29) \quad \mathbb{P}_n := \bigcup_{k \geq 0} \mathcal{P}_n^k .$$

Finally, remark that we have

$$(30) \quad \mathbb{P} = \bigcup_{n \geq 0} \mathbb{P}_n .$$

This ends the partition of our set  $\mathbb{P}$ .

## 5.2. Main Proposition; ending the proof.

In what follows, we will state the key result on which our theorem is based. The proof of this proposition will be postponed for the next sections. With the notations from the previous section, we have

**Main Proposition.** *Fix  $n \in \mathbb{N}$ . Then there exist a constant  $\eta \in (0, \frac{1}{2})$  depending only on  $d$  such that*

$$\left\| T^{\mathbb{P}_n} f \right\|_p \lesssim_{p,d} 2^{-n\eta(1-\frac{1}{p^*})} \|f\|_p ,$$

for all  $f \in L^p(\mathbb{T})$ .

If we believe this for the moment, then we trivially have

$$\|Tf\|_p \leq \sum_n \left\| T^{\mathbb{P}_n} f \right\|_p \lesssim_{p,d} \sum_n 2^{-n\eta(1-\frac{1}{p^*})} \|f\|_p \lesssim_{p,d} \|f\|_p .$$

## 6. Reduction of the main proposition

In this section, we will present the strategy needed to prove our main proposition.

**6.1. Preparatives.** We first introduce a qualitative concept that characterizes the overlapping relation between tiles.

**Definition 3.** Let  $P_j = [\alpha_j^1, \alpha_j^2, \dots, \alpha_j^d, I_j] \in \mathbb{P}$  with  $j \in \{1, 2\}$ . We say that

- $P_1 \leq P_2$  iff  $I_1 \subseteq I_2$  and  $\exists q \in P_2$  such that  $q \in P_1$ ,
- $P_1 \trianglelefteq P_2$  iff  $I_1 \subseteq I_2$  and  $\forall q \in P_2$  we have  $q \in P_1$ .

**Definition 4.** We say that a set of tiles  $\mathcal{P} \subset \mathbb{P}$  is a **tree** (relative to  $\leq$ ) with top  $P_0$  if the following conditions<sup>7</sup> are satisfied:

- 1)  $\forall P \in \mathcal{P} \Rightarrow \frac{3}{2}P \leq 10P_0$
- 2) if  $P \in \mathcal{P}$  and  $P' \in N(P)$  such that  $\frac{3}{2}P' \leq P_0$  then  $P' \in \mathcal{P}$
- 3) if  $P_1, P_2 \in \mathcal{P}$  and  $P_1 \leq P \leq P_2$  then  $P \in \mathcal{P}$ .

**Definition 5.** We say that a set of tiles  $\mathcal{P} \subset \mathbb{P}$  is a **sparse tree** if  $\mathcal{P}$  is a tree and for any  $P \in \mathcal{P}$  we have

$$(31) \quad \sum_{\substack{P' \in \mathcal{P} \\ I_{P'} \subseteq I_P}} |I_{P'}| \leq C |I_P|,$$

where here  $C > 0$  is an absolute constant.

**Definition 6.** Fix  $n \in \mathbb{N}$ . We say that  $\mathcal{P} \subseteq \mathbb{P}_n$  is an  $L^\infty$ -**forest** (of  $n^{\text{th}}$ -generation) if

- i)  $\mathcal{P}$  is a collection of separated trees, i.e.

$$\mathcal{P} = \bigcup_{j \in \mathbb{N}} \mathcal{P}_j$$

with each  $\mathcal{P}_j$  a tree with top  $P_j = [\alpha_j^1, \alpha_j^2, \dots, \alpha_j^d, I_j]$  and such that

$$(32) \quad \forall k \neq j \ \& \ \forall P \in \mathcal{P}_j \quad 2P \not\leq 10P_k.$$

- ii) the counting function

$$(33) \quad \mathcal{N}_{\mathcal{P}}(x) := \sum_j \chi_{I_j}(x)$$

obeys the estimate  $\|\mathcal{N}_{\mathcal{P}}\|_{L^\infty} \lesssim n 2^n$ .

Further on, if  $\mathcal{P} \subseteq \mathbb{P}_n$  only consists of sparse separated trees then we refer at  $\mathcal{P}$  as a **sparse  $L^\infty$ -forest**.

**Definition 7.** A set  $\mathcal{P} \subseteq \mathbb{P}_n$  is called a **BMO-forest** (of  $n^{\text{th}}$ -generation) or just simply a **forest**<sup>8</sup> if

- i)  $\mathcal{P}$  may be written as

$$\mathcal{P} = \bigcup_{j \in \mathbb{N}} \mathcal{P}_j$$

<sup>7</sup>To avoid the boundary problems arising from the definition of our tiles and from the use of a single dyadic grid, we will often involve in our reasonings a dilation factor (of the tiles).

<sup>8</sup>When the context is clear we may no longer specify the order of the generation.

with each  $\mathcal{P}_j$  an  $L^\infty$ -**forest** (of  $n^{\text{th}}$ -generation);

ii) for any  $P \in \mathcal{P}_j$  and  $P' \in \mathcal{P}_k$  with  $j, k \in \mathbb{N}$ ,  $j < k$  we either have  $I_P \cap I_{P'} = \emptyset$  or

$$|I_{P'}| \leq 2^{j-k} |I_P|.$$

As before, if  $\mathcal{P} \subseteq \mathbb{P}_n$  only consists of sparse  $L^\infty$ -forests, then, we refer at  $\mathcal{P}$  as a **sparse forest**.

**Observation.** Notice that if  $\mathcal{P} \subseteq \mathbb{P}_n$  is a forest then, due to ii) above, the counting function

$$(34) \quad \mathcal{N}_{\mathcal{P}} := \sum_j \mathcal{N}_{\mathcal{P}_j}$$

obeys the estimate

$$\|\mathcal{N}_{\mathcal{P}}\|_{BMO_C} \lesssim n 2^n$$

hence the alternative name of the  $BMO$ -forest.

In fact, based on the construction of the set  $\mathbb{P}_n$ , we actually have the stronger estimate  $\|\mathcal{N}_{\mathcal{P}}\|_{BMO_C} \lesssim 2^n$ . Moreover, notice that if  $\mathcal{P} \subseteq \mathbb{P}_n$  is a collection of separated trees then  $\mathcal{P}$  is automatically a  $(BMO)$ -forest.

Now we can state the main results of this section; their proofs will be postponed until Section 7.

**Proposition 1.** *Let  $\mathcal{P} \subseteq \mathbb{P}_n$  be a sparse forest. Then there exists  $\eta \in (0, 1/2)$ , depending only on the degree  $d$ , such that for  $1 < p < \infty$  we have*

$$\|T^{\mathcal{P}}\|_p \lesssim_{p,d} 2^{-n\eta(1-\frac{1}{p^*})}.$$

**Proposition 2.** *Let  $\mathcal{P} \subseteq \mathbb{P}_n$  be a forest. Then there exists  $\eta \in (0, 1/2)$ , depending only on the degree  $d$ , such that for  $1 < p < \infty$  we have*

$$\|T^{\mathcal{P}}\|_p \lesssim_{p,d} 2^{-n\eta(1-\frac{1}{p^*})}.$$

## 6.2. Reduction of the Main Proposition to Proposition 2.

In this section our goal is to show that, for a fixed  $n$ , the set  $\mathbb{P}_n$  can be decomposed into a controlled number of forests.

We start by reminding that, with the same notations as in Section 5, we have

$$\mathbb{P}_n = \bigcup_{k \geq 0} \mathcal{P}_n^k.$$

Our claim is that each  $\mathcal{P}_n^k$  can be decomposed in a union of at most  $cn$   $L^\infty$ -forests  $\{\mathcal{P}_n^{k,s}\}_{s \in \{1, \dots, cn\}}$ .



If we believe this for the moment, then denoting with  $\mathbb{P}_n^s := \bigcup_k \mathcal{P}_n^{k,s}$  we obviously have that  $\mathbb{P}_n^s$  is a forest. Thus, since

$$\mathbb{P}_n := \bigcup_{s=0}^{cn} \mathbb{P}_n^s,$$

we conclude that  $\mathbb{P}_n$  can be written as a union of at most  $cn$  forests.

We now return to our observation regarding the set decomposition of  $\mathcal{P}_n^k$  into a controlled number of  $L^\infty$ -forests.

Indeed, for proving this we will make use of the following facts:

- all the maximal elements in the family  $\mathcal{P}_n^k$  are gathered in the set

$$\tilde{\mathcal{P}}_n^{k,max} := \mathcal{P}_n^{k,max} \setminus \mathcal{P}_n^{k+1,max};$$

- the counting function  $\mathcal{N}_n^k$  obeys  $\|\mathcal{N}_n^k\|_{L^\infty(A_n^k \setminus A_n^{k+1})} \lesssim n 2^n$ ; thus, if we define

$\tilde{\mathcal{I}}_n^k := \{I \mid P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \tilde{\mathcal{P}}_n^{k,max}\}$  then the counting function  $\tilde{\mathcal{N}}_n^k := \sum_{I \in \tilde{\mathcal{I}}_n^k} \chi_I$  obeys  $\|\tilde{\mathcal{N}}_n^k\|_{L^\infty} \lesssim n 2^n$ .

In what follows, once that we fix  $k \in \mathbb{N}$  and the family  $\mathcal{P}_n^k$ , for notational simplicity we decide to drop the  $k$ -dependence of all the notions previously defined.

The main challenge in proving our claim is to create “spaces” (*i.e.* separation) among trees inside our family  $\mathcal{P}_n$ . But for this, we will need first to create the tree-structures. Thus, our first step is to ‘stick’ every tile  $P \in \mathcal{P}_n$  to a top (maximal tile with respect of “ $\leq$ ”). For this, we will proceed as follows:

Firstly, we notice that from the construction of our set  $\mathcal{P}_n$  we have that for any  $P \in \mathcal{P}_n$  the relation  $\frac{|E(P)|}{|I_P|} < 2^{-n+1}$  holds. Moreover, we actually have that  $A_{\mathcal{P}_n, A_n}(P) \in [2^{-n}, 2^{-n+1})$ .

Next, we select the tiles  $\{\bar{P}_j\} \subseteq \tilde{\mathcal{P}}_n^{max}$  to be those  $(d+1)$ -tuples which are maximal with respect to  $\leq$  and obey  $\frac{|E(P)|}{|I_P|} \geq 2^{-n}$ . Proceeding as in [9], we define

$$(35) \quad \tilde{\mathcal{P}}_n := \{P \in \mathcal{P}_n \mid \exists j \in \mathbb{N} \text{ s.t. } 4P \triangleleft \bar{P}_j\}$$

and set

$$\mathcal{C}_n := \left\{ P \in \mathcal{P}_n \mid \text{there are no chains } P \leq P_1 \leq \dots \leq P_n \text{ \& } \{P_j\}_{j=1}^n \subseteq \mathcal{P}_n \right\}.$$

With this done, it is easy to see that

$$\mathcal{P}_n \setminus \mathcal{C}_n \subseteq \tilde{\mathcal{P}}_n.$$

Now, defining the set  $\mathcal{D}_n \subseteq \mathcal{C}_n$  with the property  $\mathcal{P}_n \setminus \mathcal{D}_n = \tilde{\mathcal{P}}_n$ , we remark that  $\mathcal{D}_n$  breaks up as a disjoint union of at most  $n$  sets  $\mathcal{D}_{n1} \cup \mathcal{D}_{n2} \cup \dots \cup \mathcal{D}_{nn}$  with no two tiles in the same  $\mathcal{D}_{nj}$  comparable. As a consequence,  $\mathcal{D}_n$  may be written as a union of at most  $n$  sparse  $L^\infty$ -forests and hence we can erase this set from  $\mathcal{P}_n$  without affecting our claim.

Thus, in what follows, it will be enough to limit ourselves to the set of tiles  $\tilde{\mathcal{P}}_n$  which for convenience we will re-denote it with  $\mathcal{P}_n$ .

As announced, we will show that

$$\mathcal{P}_n = \bigcup_{j=1}^{cn} \mathcal{S}_{nj},$$

with each  $\mathcal{S}_{nj}$  a forest. As in [9], we set

$$B(P) := \# \{j \mid 4P \leq \bar{P}_j\} \quad \forall P \in \mathcal{P}_n,$$

and

$$\mathcal{P}_{nj} := \{P \in \mathcal{P}_n \mid 2^j \leq B(P) < 2^{j+1}\} \quad \forall j \in \{0, \dots, cn\}.$$

Fixing now a family  $\mathcal{P}_{nj}$  we look for candidates for the tops of the future trees. More exactly, take  $\{P^r\}_{r \in \{1, \dots, s\}} \subseteq \mathcal{P}_{nj}$  to be those tiles with the property that  $4P^r$  are maximal<sup>9</sup> elements with respect to the relation  $\leq$  inside the set  $4\mathcal{P}_{nj}$ .

Now, in all the reasonings that we will make further, we will use the following four essential properties:

- (A)  $4P^l \leq 4P^k \Rightarrow I_l = I_k$ ;
- (B)  $\forall P \in \mathcal{P}_{nj} \exists P^l$  s.t.  $4P \leq 4P^l$ ;
- (C) If  $P \in \mathcal{P}_{nj}$  s.t.  $\exists k \neq l$  with  $\left\{ \begin{smallmatrix} 4P \leq 4P^l \\ 4P \leq 4P^k \end{smallmatrix} \right\}$ , then  $\left\{ \begin{smallmatrix} 4P^k \leq 4P^l \\ 4P^l \leq 4P^k \end{smallmatrix} \right\}$ ;
- (D) If  $P_j = [\alpha_j^1, \alpha_j^2, \dots, \alpha_j^d, I_j] \in \mathbb{P}$  with  $j \in \{1, 2\}$  s.t.  $|I_1| \neq |I_2|$ , then  $|I_1| \leq 2^{-D} |I_2|$  or  $|I_2| \leq 2^{-D} |I_1|$ .

(While (A), (B) and (C) are derived from the definition of  $\mathcal{P}_{nj}$  and the way in which we have chosen  $\{P^r\}_{r \in \{1, \dots, s\}}$ , property (D) follows from the assumption made at the end of Section 3.)

The next step, is to discard the tiles that are not “close” to our new maximal elements (recall that our final goal is to construct separated trees). For technical reasons, we also get rid of the maximal elements together with their neighbors and respectively of the minimal elements.

More exactly, we define:

$$\mathcal{R}_{nj} := \left\{ P \in \mathcal{P}_{nj} \mid \forall P^l \Rightarrow \frac{3}{2}P \not\leq P^l \right\} \cup$$

$$\left\{ P \in \mathcal{P}_{nj} \mid \exists l \text{ st } |I_P| = |I_{P^l}|, \frac{3}{2}P \leq P^l \right\} \cup \left\{ P \in \mathcal{P}_{nj} \mid P \text{ minimal} \right\},$$

and we set

$$\mathcal{S}_{nj} := \mathcal{P}_{nj} \setminus \mathcal{R}_{nj}.$$

---

<sup>9</sup>Here we use the following convention: let be  $\mathcal{D}$  a collection of tiles;  $P$  is maximal (relative to  $\leq$ ) in  $\mathcal{D}$  iff  $\forall P' \in \mathcal{D}$  such that  $P \leq P'$  we also have  $P' \leq P$ .

Now, using the properties (B) and (D), it is easy to see that  $\mathcal{R}_{nj}$  forms a **negligible**<sup>10</sup> set of tiles.

For the remaining set  $\mathcal{S}_{nj}$ , one should follow the steps below (here we use the properties (A)-(D)):

- Set  $S_k = \{P \in \mathcal{S}_{nj} \mid \frac{3}{2}P \leq P^k\}$ ; without loss of generality we may suppose  $\mathcal{S}_{nj} = \bigcup_{k=1}^s S_k$ ;
- Introduce the following relation among the sets  $\{S_k\}_k$ :

$$S_k \propto S_l$$

if  $\exists P_1 \in S_k$  and  $\exists P_2 \in S_l$  such that  $2P_1 \leq 10P^l$  or  $2P_2 \leq 10P^k$ ;

- Deduce that “ $\propto$ ” becomes an order relation and that

$$S_k \propto S_l \Rightarrow 4P^k \leq 4P^l \Rightarrow I^k = I^l;$$

- Let  $\hat{k} := \{m \mid S_m \propto S_k\}$ ; then the cardinality of  $\hat{k}$  is at most  $c(d)$ , and for

$$\hat{S}_k := \bigcup_{m \in \hat{k}} S_m,$$

one has that  $\hat{S}_k$  is a tree having as a top any  $P^l$  with  $l \in \hat{k}$ .

- Consider  $\{\hat{S}_k\}_k$  such that in this enumeration all the elements are distinct. Conclude then that the relation “ $\propto$ ”, can be meaningfully extended<sup>11</sup> among the sets  $\{\hat{S}_k\}_k$  and that

$$\hat{S}_k \propto \hat{S}_l \Rightarrow k = l.$$

Consequently, from the algorithm just described, we deduce that the set

$$\mathcal{S}_{nj} = \bigcup_k \hat{S}_k$$

is a forest as in Definition 7.

## 7. Some technicalities - the proofs of Propositions 1 and 2

### 7.1. Proof of Proposition 1.

We begin by restating the result that we need to prove:

**Proposition 1.** *Let  $\mathcal{P} \subseteq \mathbb{P}_n$  be a sparse forest. Then there exists  $\eta \in (0, 1/2)$ , depending only on the degree  $d$ , such that for  $1 < p < \infty$  we have*

$$\|T^{\mathcal{P}}\|_p \lesssim_{p,d} 2^{-n\eta(1-\frac{1}{p^*})}.$$

<sup>10</sup>The set can be written as a union of at most  $c(d)$  families of incomparable tiles.

<sup>11</sup>In this case, the role played by the maximal element  $P^k$  in the initial definition is now taken by the top of the corresponding tree.

### 7.1.1. The $L^2$ bound.

Starting as in the corresponding proof of Proposition 1 in [9], we have

$$\begin{aligned}
\int_{\mathbb{T}} \left| (T^P)^* f(x) \right|^2 dx &\lesssim \left| \sum_{\substack{P' \in \mathcal{P} \\ P' = [\alpha', I']}} \int_{\mathbb{T}} f(x) \left\{ \sum_{\substack{P = [\alpha, I] \in \mathcal{P} \\ |I| \leq |I'|}} \overline{T_{P'} T_P^* f(x)} \right\} dx \right| \\
&\lesssim \sum_{P' \in \mathcal{P}} \int_{E(P')} |f| \left\{ \sum_{P \in a(P')} [\Delta(P, P')]^{1/d} \frac{\int_{E(P)} |f|}{|I'|} \right\} \\
&+ \sum_{P' \in \mathcal{P}} \int_{E(P')} |f| \left\{ \sum_{P \in b(P')} [\Delta(P, P')]^{1/d} \frac{\int_{E(P)} |f|}{|I'|} \right\} =_{def} A + B
\end{aligned}$$

where for the third inequality we used the estimate (cf. Lemma 0)

$$|T_{P'} T_P^* f(x)| \lesssim [\Delta(P, P')]^{1/d} \frac{\int_{E(P)} |f|}{|I'|} \chi_{E(P')}(x)$$

together with the following notations:

$$a(P') = \{P = [\alpha, I] \in \mathcal{P}, |I| \leq |I'| \text{ \& } I^* \cap I'^* \neq \emptyset \mid \Delta(P, P') \leq 2^{2n\epsilon}\},$$

$$b(P') = \{P = [\alpha, I] \in \mathcal{P}, |I| \leq |I'| \text{ \& } I^* \cap I'^* \neq \emptyset \mid \Delta(P, P') \geq 2^{2n\epsilon}\}.$$

(Here  $\epsilon \in (0, 1)$  is some constant that may be chosen later.)

Further, we have

$$A \lesssim \sum_{P' \in \mathcal{P}} \int_{E(P')} |f(x)| \left\{ \frac{1}{|I'|} \sum_{P \in a(P')} \int_{E(P)} |f| \right\} dx = \int |f| V_a(|f|),$$

where by definition

$$(36) \quad V_a(f) := \sum_{P' \in \mathcal{P}} \frac{\chi_{E(P')}}{|I_{P'}|} \sum_{P \in a(P')} \int_{E(P)} f,$$

and similarly

$$B \lesssim \sum_{P' \in \mathcal{P}} \int_{E(P')} |f(x)| \left\{ \frac{2^{-n \frac{\epsilon}{d}}}{|I'|} \sum_{P \in b(P')} \int_{E(P)} |f| \right\} dx = 2^{-n \frac{\epsilon}{d}} \int |f| V_b(|f|),$$

where by definition

$$(37) \quad V_b(f) := \sum_{P' \in \mathcal{P}} \frac{\chi_{E(P')}}{|I_{P'}|} \sum_{P \in b(P')} \int_{E(P)} f.$$

We will now focuss on providing  $L^2$ -bounds on  $V_a(f)$ . Let  $1 < r < 2$  and suppose wlog that  $f \geq 0$ ; then

$$V_a(f) \leq \sum_{P' \in \mathcal{P}} \chi_{E(P')} \left( \frac{\int_{I_{P'}^*} f^r}{|I_{P'}|} \right)^{\frac{1}{r}} \frac{\|\sum_{P \in a(P')} \chi_{E(P)}\|_{r'}}{|I_{P'}|^{\frac{1}{r'}}}.$$

The first key observation derived from the structure of the set  $\mathcal{P}$  and the definition of  $a(P)$  is that

$$(38) \quad \left\| \sum_{P \in a(P')} \chi_{E(P)} \right\|_{r'} \lesssim 2^{-n(\frac{1}{r'} - 100\epsilon)} |I_{P'}|^{\frac{1}{r'}}.$$

Indeed, one first notice that for any collection  $\mathcal{A} \subseteq \mathcal{P}$  of *incomparable* tiles one has (see [9])

$$\left\| \sum_{\substack{P \in a(P') \\ P \in \mathcal{A}}} \chi_{E(P)} \right\|_1 \lesssim 2^{-n(1-10\epsilon)} |I_{P'}|.$$

On the other hand by the definition of  $\mathcal{A}$  we do have that

$$\left\| \sum_{P \in \mathcal{A}} \chi_{E(P)} \right\|_{\infty} \leq 1.$$

By interpolation we deduce that (38) holds for  $\mathcal{P} \cap a(P')$  set of incomparable tiles.

Now, for the general case, due to *ii*) in Definition 7, it is enough to show (38) for  $\mathcal{P}$  a sparse  $L^\infty$ -forest. Then, by Definition 6, we have that  $\mathcal{P} \cap a(P') = \bigcup_j \mathcal{P}_j$  with  $\{\mathcal{P}_j\}_j$  sparse separated trees. Further set  $\text{top} \mathcal{P}_j = P_j$  and let

$$\mathcal{P}_j^1 = \{P \in \mathcal{P}_j \mid \text{there is no chain } P < P^1 < \dots < P^n = P_j \text{ s.t. } P^k \in \mathcal{P}_j\},$$

and

$$\mathcal{P}_j^2 := \mathcal{P}_j \setminus \mathcal{P}_j^1.$$

In the above setting, we notice that  $\mathcal{P} \cap a(P')$  can be written as

$$\left( \bigcup_{k=1}^n \mathcal{A}_k \right) \cup \left( \bigcup_j \mathcal{P}_j^2 \right),$$

with each  $\mathcal{A}_k$  set of incomparable tiles and

$$\sum_{j, \mathcal{P}_j^2 \neq \emptyset} \chi_{I_{P_j}} \leq 1.$$

Finally, we already know that (38) applies for each  $\mathcal{A}_k$  while from the fact that each  $\mathcal{P}_j$  is a sparse tree we deduce that

$$\left\| \sum_{P \in \mathcal{P}_j^2} \chi_{E(P)} \right\|_{r'} \lesssim 2^{-n \frac{1}{r'}} |I_{P_j}|^{\frac{1}{r'}}.$$

Combining all this facts together we conclude that (38) holds in full generality.

Now, in order to control the term  $A$  it remains to show that for

$$\mathcal{V}f := \sum_{P \in \mathcal{P}} \chi_{E(P)} \left( \frac{\int_{I_P^*} f^r}{|I_P|} \right)^{\frac{1}{r}},$$

we have

$$(39) \quad \|\mathcal{V}f\|_2 \lesssim_r \|f\|_2.$$

Set now  $\mathcal{I} := \{I \mid \exists P \in \mathcal{P} \text{ s.t. } I_P = I\}$  and  $E(I) := \bigcup_{I_P=I} E(P)$ . Rewrite  $\mathcal{V}$  as follows:

$$\mathcal{V}f = \sum_{I \in \mathcal{I}} \chi_{E(I)} \left( \frac{\int_{I^*} f^r}{|I|} \right)^{\frac{1}{r}}.$$

Denote with  $\mathcal{I}_m := \{I \in \mathcal{I} \mid \frac{\int_{I^*} f^r}{|I|} \approx 2^m\}$  and notice that  $\mathcal{I} = \bigcup_{m \in \mathbb{Z}} \mathcal{I}_m$ . Also denote with  $\mathcal{I}_m^{max}$  the set of maximal intervals (with respect of inclusion) in  $\mathcal{I}_m$ . For each  $m$ , notice then that  $\mathcal{I}_m^{max}$  consists of pairwise disjoint intervals.

Then we have

$$\mathcal{V}f = \sum_{m \in \mathbb{Z}} \sum_{I \in \mathcal{I}_m} \chi_{E(I)} \left( \frac{\int_{I^*} f^r}{|I|} \right)^{\frac{1}{r}} \approx \sum_{m \in \mathbb{Z}} \sum_{J \in \mathcal{I}_m^{max}} \sum_{\substack{I \subseteq J \\ I \in \mathcal{I}_m}} 2^{\frac{m}{r}} \chi_{E(I)},$$

and thus

$$\begin{aligned} \|\mathcal{V}f\|_2^2 &\approx \sum_{m, m'} \sum_{\substack{J \in \mathcal{I}_m^{max} \\ J' \in \mathcal{I}_{m'}^{max}}} 2^{\frac{m+m'}{r}} \int \left( \sum_{\substack{I \subseteq J \\ I \in \mathcal{I}_m}} \chi_{E(I)} \right) \left( \sum_{\substack{I' \subseteq J' \\ I' \in \mathcal{I}_{m'}}} \chi_{E(I')} \right) \\ &\approx \sum_m \sum_{m' \geq m} \sum_{J \in \mathcal{I}_m^{max}} \sum_{\substack{J' \subseteq J \\ J' \in \mathcal{I}_{m'}^{max}}} 2^{\frac{m+m'}{r}} \int \left( \sum_{\substack{I \subseteq J \\ I \in \mathcal{I}_m}} \chi_{E(I)} \right) \left( \sum_{\substack{I' \subseteq J' \\ I' \in \mathcal{I}_{m'}}} \chi_{E(I')} \right) \end{aligned}$$

Using now the Carleson packing condition (with  $1 \leq q < \infty$  and  $J \subseteq [0, 1]$ )

$$(40) \quad \left\| \sum_{I \subseteq J} \chi_{E(I)} \right\|_q^q \lesssim |J|,$$

and applying Cauchy-Schwarz for  $1 < p < r < 2$ , we further have

$$\begin{aligned} \|\mathcal{V}f\|_2^2 &\lesssim \sum_m \sum_{m' \geq m} 2^{\frac{m+m'}{r}} \sum_{J \in \mathcal{I}_m^{max}} \left\| \sum_{I \subseteq J} \chi_{E(I)} \right\|_{p'} \left\| \sum_{\substack{J' \subseteq J \\ J' \in \mathcal{I}_{m'}^{max}}} \sum_{I' \subseteq J'} \chi_{E(I')} \right\|_p \\ &\lesssim \sum_m \sum_{J \in \mathcal{I}_m^{max}} 2^{\frac{m}{r}} |J|^{\frac{1}{p'}} \sum_{m' \geq m} 2^{\frac{m'}{r}} \left( \sum_{\substack{J' \subseteq J \\ J' \in \mathcal{I}_{m'}^{max}}} |J'| \right)^{\frac{1}{p}} \\ &\lesssim \sum_m \sum_{J \in \mathcal{I}_m^{max}} 2^{\frac{m}{r}} |J|^{\frac{1}{p'}} \sum_{m' \geq m} 2^{\frac{m'}{r}} 2^{-\frac{m'}{p}} \left( \int_J f^r \right)^{\frac{1}{p}} \end{aligned}$$

$$\lesssim \sum_m \sum_{J \in \mathcal{I}_m^{max}} 2^{\frac{2m}{r}} |J| \lesssim \sum_m 2^{\frac{2m}{r}} 2^{-m} \int_{(M_r f)^r \gtrsim 2^m} (M_r f)^r \lesssim \int (M_r f)^2 \lesssim_r \int f^2,$$

where here we denoted  $M_r f(x) := \left( \sup_{x \in I} \frac{\int_I |f|^r}{|I|} \right)^{\frac{1}{r}}$ .

Thus, combining now (38) and (39), we conclude that

$$A \lesssim_r 2^{-n(\frac{1}{r'} - 100\epsilon)} \|f\|_2.$$

The  $B$  term can be similarly treated if one replaces (38) with just

$$(41) \quad \left\| \sum_{P \in b(P')} \chi_{E(P)} \right\|_{r'} \lesssim |I_{P'}|^{\frac{1}{r'}},$$

thus obtaining

$$B \lesssim 2^{-n \frac{\epsilon}{d}} \|f\|_2.$$

Now, properly choosing  $r$  and  $\epsilon$ , we conclude that there exists  $\eta = \eta(d) \in (0, 1)$  such that

$$\|T^{\mathcal{P}}\|_2 \lesssim 2^{-n \frac{\eta}{2}}.$$

This ends our proof.

**7.1.2. The  $L^p$  bound.** Suppose first that  $2 < p < \infty$ ; we then observe that on one hand

$$|T^{\mathcal{P}} f| \leq \sum_{P \in \mathcal{P}} |T^P f| \lesssim \sum_{P \in \mathcal{P}} \frac{\int_{I_{P^*}} |f|}{|I_P|} \chi_{E(P)} \lesssim \|f\|_{\infty} \sum_{P \in \mathcal{P}} \chi_{E(P)}.$$

On the other hand

$$\left\| \sum_{P \in \mathcal{P}} \chi_{E(P)} \right\|_{L_{\exp}} \leq \left\| \sum_{P \in \mathcal{P}} \chi_{E(P)} \right\|_{BMO_D} \lesssim 1.$$

From this we conclude that

$$\|T^{\mathcal{P}}\|_{\infty \rightarrow L_{\exp}} \lesssim 1.$$

Interpolating now between  $L^2 \rightarrow L^2$  and  $L^{\infty} \rightarrow L_{\exp}$  we obtain the desired conclusion.

For the case  $1 < p < 2$  we need to focus on the behavior of  $T^{\mathcal{P}*}$ . Indeed, on the one hand we know that

$$\|T^{\mathcal{P}*}\|_{2 \rightarrow 2} = \|T^{\mathcal{P}}\|_{2 \rightarrow 2} \lesssim 2^{-n \frac{\eta}{2}}.$$

On the other hand, for  $f \in L^{\infty}$  we have

$$|T^{\mathcal{P}*} f| \leq \sum_{P \in \mathcal{P}} |T^{P*} f| \lesssim \sum_{P \in \mathcal{P}} \frac{\int_{E(P)} |f|}{|I_P|} \chi_{I_{P^*}} \lesssim \|f\|_{\infty} \sum_{P \in \mathcal{P}} \frac{|E(P)|}{|I_P|} \chi_{I_{P^*}}.$$

Thus

$$\|T^{\mathcal{P}*} f\|_{L_{\exp}} \leq \|f\|_{\infty} \left\| \sum_{P \in \mathcal{P}} \frac{|E(P)|}{|I_P|} \chi_{I_{P^*}} \right\|_{BMO_D} \lesssim \|f\|_{\infty}$$

from which we conclude that

$$\left\| T^{\mathcal{P}*} \right\|_{\infty \rightarrow L_{\exp}} \lesssim 1.$$

The claim now follows by interpolation.<sup>12</sup>

□

## 7.2. Preparatives for the proof of Proposition 2.

As the name suggest, this section is meant for “preparing the ground” for the proof of Proposition 2. Most of the results presented here, have a direct analogue in either [4] or [9], and thus, we will not insist on their proofs but only treat the sensitive points that are different.

### 7.2.1. $L^2$ -results. Main Lemma.

We start this section with the tree-estimate given by

**Lemma 1.** *Let  $\delta > 0$  be fixed and, let  $\mathcal{P} \subseteq \mathbb{P}$  be a tree such that*

$$A_0(P) < \delta \quad \forall \quad P \in \mathcal{P}.$$

*Then*

$$(42) \quad \left\| T^{\mathcal{P}} \right\|_2 \lesssim_d \delta^{\frac{1}{2}}.$$

**Definition 8.** *Fix a number  $\delta \in (0, 1)$ . Let  $\mathcal{P}_1$  and  $\mathcal{P}_2$  be two trees with tops  $P_1 = [\alpha_1^1, \alpha_1^2, \dots, \alpha_1^d, I_1]$  and respectively  $P_2 = [\alpha_2^1, \alpha_2^2, \dots, \alpha_2^d, I_2]$ ; we say that  $\mathcal{P}_1$  and  $\mathcal{P}_2$  are  $(\delta^{-1})$ -separated if either  $I_1 \cap I_2 = \emptyset$  or else*

- i)  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}_1$  &  $I \subseteq I_2 \Rightarrow [\Delta(P, P_2)] < \delta$ ,*
- ii)  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}_2$  &  $I \subseteq I_1 \Rightarrow [\Delta(P, P_1)] < \delta$ .*

**Notation:** Let  $\mathcal{P}_1$  and  $\mathcal{P}_2$  be two trees as in Definition 8. Take  $q_j$  to be the central polynomial of  $P_j$  ( $j \in \{1, 2\}$ ), set  $q_{1,2} = q_1 - q_2$ , and then define

- $I_s$  - the **separation set** (relative to the intersection) of  $\mathcal{P}_1$  and  $\mathcal{P}_2$  by

$$I_s = \mathcal{I}_s \left( \eta_{1,2}, c(d)\delta^{-1}, q_{1,2}, \tilde{I}_1 \cap \tilde{I}_2 \right),$$

where  $\eta_{1,2} = \eta(\tilde{I}_1 \cap \tilde{I}_2)$ ;

- $I_c$  - the **( $\epsilon$ -)critical intersection set** (between  $\mathcal{P}_1$  and  $\mathcal{P}_2$ ) by

$$I_c = \bigcup_{j=1}^r \mathcal{I}_c \left( \eta(I_s^j), w(I_s^j), q_{1,2}, I_s^j \right),$$

where  $\epsilon$  is some small fixed positive real number and  $I_s = \bigcup_{j=1}^r I_s^j$  is the decomposition of  $I_s$  into maximal disjoint intervals  $\{I_s^j\}_{j \in \{1, \dots, r\}}$  with  $r \leq 2d$ .

---

<sup>12</sup>We use here the fact that  $\left\| T^{\mathcal{P}*} \right\|_{p' \rightarrow p'} = \left\| T^{\mathcal{P}} \right\|_{p \rightarrow p}$ .



**Observation 5.** It is important to notice the following three properties of our above-defined sets; these facilitate the adaptation of the reasonings involved in the proofs of Lemmas 2 and 4 to those of the corresponding lemmas in [9]:

- 1) for all dyadic  $J \subset \tilde{I}_1 \cap \tilde{I}_2$  such that  $I_s \cap 5J = \emptyset$  we have  $(c(d) \leq d^d)$

$$\inf_{x \in J} |q_{1,2}(x)| \leq \sup_{x \in J} |q_{1,2}(x)| \leq c(d) \inf_{x \in J} |q_{1,2}(x)|.$$

- 2)  $\forall P \in \mathcal{P}_1 \cup \mathcal{P}_2$  and  $j \in \{1, \dots, r\}$  if  $I_s^j \cap 5\tilde{I}_P \neq \emptyset$  then  $|I_P| > |I_s^j|$ .

- 3)  $\forall P \in \mathcal{P}_1 \cup \mathcal{P}_2$  we have (for  $\epsilon$  properly chosen)  $|\tilde{I}_P \cap I_c| < \delta^{\frac{1}{100d}} |I_P|$ .

Indeed, these facts are an easy consequence of the results mentioned in the Appendix and the way in which  $I_s$  and  $I_c$  are defined.

(Remark that property 1) above implies the following relation:

$$\forall P \in \mathcal{P}_1 \text{ such that } I_s \cap 5\tilde{I}_P = \emptyset \text{ and } I_P \subset I_2 \text{ we have}$$

$$\text{Graph}(q_2) \cap (c(d)\delta^{-1}) \hat{P} = \emptyset.$$

Of course, the same is true for the symmetric relation, *i.e.* replacing the index 1 with 2 and vice versa.)

**Lemma 2.** Let  $\{\mathcal{P}_j\}_{j \in \{1,2\}}$  be two  $\delta^{-1}$ -separated trees with tops

$P_j = [\alpha_j^1, \alpha_j^2, \dots, \alpha_j^d, I_0]$ . Then, for any  $f, g \in L^2(\mathbb{T})$  and  $n \in \mathbb{N}$ , we have that

(43)

$$\left| \left\langle T^{\mathcal{P}_1^*} f, T^{\mathcal{P}_2^*} g \right\rangle \right| \lesssim_{n,d} \delta^n \|f\|_{L^2(\tilde{I}_0)} \|g\|_{L^2(\tilde{I}_0)} + \left\| \chi_{I_c} T^{\mathcal{P}_1^*} f \right\|_2 \left\| \chi_{I_c} T^{\mathcal{P}_2^*} g \right\|_2.$$

*Proof.* In what follows we intend to adapt the methods described in the proof of Lemma 2 of [9] to our context. For this, we need first to modify the definition of the sets  $\{A_l\}_l$ ; more exactly we follow the procedure below:

Let  $I_s = \bigcup_{j=1}^r I_s^j$  be the decomposition of  $I_s$  into maximal disjoint intervals ( $r \leq 2d$ ); without loss of generality we may suppose<sup>13</sup> that  $\{I_s^j\}_j$  are placed

in consecutive order with  $I_s^{j+1}$  located to the right of  $I_s^j$ . For a fixed  $j \in \{0, \dots, r\}$ , let  $W_j$  be the standard Whitney decomposition of the set  $[0, 1] \cap (I_s^j)^c \cap (I_s^{j+1})^c$  with respect to the set  $I_s^j \cup I_s^{j+1}$ ; we take  $\tilde{W}_j$  to be the “large scale” version of  $W_j$ , which is obtained as follows:

- take the union of all the intervals in  $W_j$  of length strictly smaller than  $c(d)|I_s^j|$  that approach  $I_s^j$  and denote it by  $R_j$  (we can do this in such a way that  $R_j$  can be written as a union of at most two dyadic intervals, each one of length  $c(d)|I_s^j|$ );
- apply the same procedure to obtain  $R_{j+1}$ ;
- the rest of the intervals belonging to  $W_j$  remain unchanged and are transferred to  $\tilde{W}_j$ .

Define  $\mathcal{W}_j := \tilde{W}_j \cup I_s^j \cup I_s^{j+1}$  and observe that this is a partition of  $[0, 1]$ .

<sup>13</sup>Here  $I_s^0 = I_s^{r+1} = \emptyset$

Finally, we take  $\mathcal{W}$  to be the common refinement of the partitions  $\mathcal{W}_j$ ,  $j \in \{0, \dots, r\}$ . Take now  $A_0 = \bigcup_{\substack{I \in \mathcal{W} \\ \bar{I} \cap \bar{I}_s \neq \emptyset}} I$  and set

$$\mathcal{W} = A_0 \cup \bigcup_{l=1}^k A_l.$$

Then, for  $l \in \{1, \dots, k\}$  and  $m \in \{1, 2\}$  we define the sets

$$S_{m,l} := \left\{ P \in \mathcal{P}_m \mid I_P \subset A_l \text{ \& } |I_P| \leq \frac{|A_l|}{20} \right\}.$$

Also, we take  $S_{m,0} := \mathcal{P}_m \setminus \left( \bigcup_{l=1}^k S_{m,l} \right)$ .

With this done, for  $l \in \{0, \dots, k\}$ , we set

$$T_{m,l}^* = \sum_{P \in S_{m,l}} T_P^*$$

and deduce that

$$\langle T^{\mathcal{P}_1^*}, T^{\mathcal{P}_2^*} \rangle = \sum_{n,l=0}^k \langle T_{1,l}^*, T_{2,n}^* \rangle.$$

Now, as intended, we may follow the same steps as in [9], Lemma 2 and show that

$$(44) \quad \sum_{l=0}^k \sum_{n=1}^k |\langle T_{1,l}^* f, T_{2,n}^* g \rangle| \lesssim_{n,d} \delta^n \|f\|_{L^2(\tilde{I}_0)} \|g\|_{L^2(\tilde{I}_0)};$$

$$(45) \quad |\langle T_{1,0}^* f, T_{2,0}^* g \rangle| \lesssim_{n,d} \delta^n \|f\|_{L^2(\tilde{I}_0)} \|g\|_{L^2(\tilde{I}_0)} + \left\| \chi_{I_c} T^{\mathcal{P}_1^*} f \right\|_2 \left\| \chi_{I_c} T^{\mathcal{P}_2^*} g \right\|_2.$$

finishing our proof.  $\square$

**Definition 9.** A tree  $\mathcal{P}$  with top  $P_0 = [\alpha_0^1, \alpha_0^2, \dots, \alpha_0^d, I_0]$  is called **normal** if for any  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P}$  we have  $20I \cap (I_0)^c = \emptyset$ .

**Observation 6.** Notice that if  $\mathcal{P}$  is a normal tree as above then

$$\text{supp } T^{\mathcal{P}^*} f \subseteq I_0.$$

**Definition 10.** A row is a collection  $\mathcal{P} = \bigcup_{j \in \mathbb{N}} \mathcal{P}^j$  of normal trees  $\mathcal{P}^j$  with tops  $P_j = [\alpha_j^1, \alpha_j^2, \dots, \alpha_j^d, I_j]$  such that the  $\{I_j\}$  are pairwise disjoint.

The proofs of the next two lemmas require no significant modifications from the corresponding proofs in [9].

**Lemma 3.** Let  $\mathcal{P}$  be a row as above, let  $\mathcal{P}'$  be a tree with top  $P' = [\alpha_0^{1'}, \alpha_0^{2'}, \dots, \alpha_0^{d'}, I_0']$  and suppose that  $\forall j \in \mathbb{N}$ ,  $I_0^j \subseteq I_0'$  and  $\mathcal{P}^j, \mathcal{P}'$  are  $\delta^{-1}$  separated trees; denote by  $I_c^j$  the critical intersection set between each  $\mathcal{P}^j$  and  $\mathcal{P}'$ .

Then for any  $f, g \in L^2(\mathbb{T})$  and  $n \in \mathbb{N}$  we have that

$$\left| \langle T^{\mathcal{P}'*} f, T^{\mathcal{P}*} g \rangle \right| \lesssim_{n,d} \delta^n \|f\|_2 \|g\|_2 + \left\| \sum_j \chi_{I_c^j} T^{\mathcal{P}'*} f \right\|_2 \left\| \sum_j \chi_{I_c^j} T^{\mathcal{P}*} g \right\|_2.$$

**Lemma 4.** Let  $\mathcal{P}$  be a tree with top  $P_0 = [\alpha_0^1, \alpha_0^2, \dots, \alpha_0^d, I_0]$ ; suppose also that we have a set  $A \subseteq \tilde{I}_0$  with the property that

$$(46) \quad \exists \delta \in (0, 1) \text{ st } \forall P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P} \text{ we have } |I^* \cap A| \leq \delta |I|.$$

Then  $\forall f \in L^2(\mathbb{T})$  we have

$$(47) \quad \left\| \chi_A T^{\mathcal{P}*} f \right\|_2 \lesssim \delta^{\frac{1}{2}} \|f\|_2.$$

We are now in the position to state the main result of this section.

**Main Lemma.** Let  $\mathcal{P} \subset \mathbb{P}_n$  be an  $L^\infty$ -forest (of  $n^{\text{th}}$ -generation).

Then there exists  $\eta = \eta(d) \in (0, 1)$  such that

$$\|T^{\mathcal{P}} f\|_2 \lesssim 2^{-\frac{n}{2}\eta} \|f\|_2.$$

Moreover, if  $\mathcal{P}$  is normal and  $2^{100nd}$ -separated,<sup>14</sup> then writing  $\mathcal{P}$  as a union of rows  $\{\mathcal{R}_j\}$ , we have the almost orthogonality relation

$$(48) \quad \|T^{\mathcal{P}} f\|_2^2 \lesssim \sum_j \|T^{\mathcal{R}_j} f\|_2^2$$

and hence

$$\|T^{\mathcal{P}} f\|_2 \lesssim 2^{-\frac{n}{2}} \|f\|_2.$$

*Proof.* We start by writing  $\mathcal{P} = \bigcup_{j=1}^{c n^{2n}} \mathcal{R}_j$  with  $\mathcal{R}_j$  collection of spatially disjoint trees (i.e. the time intervals of the tops are disjoint). Further, decompose each  $\mathcal{R}_j$  in a disjoint union of trees  $\{\mathcal{T}_{j,k}\}_k$ . If  $I_{jk}$  stands for the time-interval of the top of  $\mathcal{T}_{j,k}$ , define the boundary component  $\mathcal{T}_{j,k}^{bd} := \{P \in \mathcal{T}_{j,k} \mid 20I_P \cap (I_{jk})^c \neq \emptyset\}$ . Also let  $\tilde{\mathcal{T}}_{j,k}$  be the set of tiles of the smallest  $100nd$  scales in  $\mathcal{T}_{j,k}$ . Then, as one can easily notice, the set  $\bigcup_{j,k} \mathcal{T}_{j,k}^{bd} \cup \tilde{\mathcal{T}}_{j,k}$  can be decomposed in at most  $cn$  negligible sets for which we can apply Proposition 1. Thus, we can erase this set of tiles from our initial forest.

We are now left with the case  $\mathcal{P}$  normal and  $2^{100nd}$ -separated  $L^\infty$ -forest. It then remains to show, that in this new context, we have that the operators  $\{T^{\mathcal{R}_j}\}_j$  are almost orthogonal. More precisely, for  $k \neq j$ , we claim that

- i)  $\|T^{\mathcal{R}_k*} T^{\mathcal{R}_j}\|_{2 \rightarrow 2} = 0$ ;
- ii)  $\|T^{\mathcal{R}_k} T^{\mathcal{R}_j*}\|_{2 \rightarrow 2} \lesssim 2^{-5n}$ .

The first claim is a direct consequence of the pairwise disjointness of the sets  $\{\text{supp } T^{\mathcal{R}_j}\}_j$ . For the second claim one needs to make use of the strong  $(2^{100nd})$ -separateness hypothesis and successively apply Lemmas 3 and 4. We leave these details to the reader.  $\square$

<sup>14</sup>As expected, an  $L^\infty$ -forest  $\mathcal{P}$  is called normal if all the trees inside are normal; same principle applies for the  $\delta^{-1}$ -separateness condition.

### 7.2.2. $L^p$ -results.

In this (sub)section, we will only state some  $L^p$  versions of the results presented in the previous (sub)section. Their proofs follow the same lines as those described for obtaining the  $L^2$ -results at which one may add some standard interpolation techniques. Thus, we only sketch a proof for the first statement, leaving the proof details of the other lemmas to the reader.

**Lemma 5.** *Let  $\delta > 0$  be fixed and, let  $\mathcal{P} \subseteq \mathbb{P}$  be a tree such that*

$$A_0(P) < \delta \quad \forall \quad P \in \mathcal{P}.$$

*Then, for  $1 < p < \infty$ , we have*

$$(49) \quad \|T^{\mathcal{P}}\|_p \lesssim_{p,d} \delta^{\frac{1}{p}}.$$

*Proof.* We start by setting the parameters of our tree; more precisely, we fix the top  $P_0 = [\alpha_0^1, \alpha_0^2, \dots, \alpha_0^d, I_0]$ , and frequency polynomial  $q_0$ . Since we have specified the polynomial  $q_0$  we also know the form of  $Q_0$ ; suppose now that

$$Q_0(y) = \sum_{j=1}^d a_j^0 y^j.$$

Then, denoting  $g(x) = M_{1,a_1^0}^* \dots M_{d,a_d^0}^* f(x)$  and reasoning<sup>15</sup> as in the proof of Lemma 1 in [9], one can show that:

$$|T^{\mathcal{P}} f(x)| \lesssim_d M_\delta(R * g)(x) + M_\delta g(x),$$

where we set  $R(y) = \sum_{k \in \mathbb{N}D} \psi_k(y)$  (here without loss of generality we suppose that  $\mathcal{P} \subset \bigcup_{k \in \mathbb{N}} \mathbb{P}_{kD}$ ).

Now, taking into account the fact that  $\|M_\delta g\|_p \lesssim_p \delta^{\frac{1}{p}} \|g\|_p$  and  $\|R * g\|_p \lesssim_p \|g\|_p$ , we conclude that (49) holds.  $\square$

**Lemma 6.** *Let  $\mathcal{P}$  be a tree. Define the following partition of  $[0, 1]$ :*

$$\mathcal{J}_{\mathcal{P}} := \{J \text{ dyadic interv.} \subseteq [0, 1] \mid J \text{ maximal s.t. } P \in \mathcal{P} \& J \cap I_P \neq \emptyset \Rightarrow I_P \supseteq J\}.$$

*Let  $A \subseteq [0, 1]$ . For  $J \in \mathcal{J}_{\mathcal{P}}$  define  $E_A(J) := \bigcup_{P \in \mathcal{P}} (A \cap E(P) \cap J)$ .*

*Then, for  $1 < p < \infty$ , we have*

$$(50) \quad \|\chi_A T^{\mathcal{P}} f\|_p \lesssim_p \left( \sup_{J \in \mathcal{J}_{\mathcal{P}}} \frac{|E_A(J)|}{|J|} \right)^{\frac{1}{p}} \|f\|_p.$$

**Lemma 7.** *Let  $\mathcal{P}$  be a tree. Suppose also that we have a set  $A \subseteq [0, 1]$  with the property that*

$$(51) \quad \exists \delta \in (0, 1) \text{ st } \quad \forall P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathcal{P} \text{ we have } |I^* \cap A| \leq \delta |I|.$$

---

<sup>15</sup>Here the key element is the following perspective: “A tree behaves like a maximal truncated Hilbert transform”.

Then for any  $f \in L^p(\mathbb{T})$  we have

$$(52) \quad \left\| \chi_A T^{\mathcal{P}^*} f \right\|_p \lesssim_p \delta^{\frac{1}{p}} \sup_{P \in \mathcal{P}} A(P)^{\frac{1}{p^*}} \|f\|_p.$$

### 7.3. Proof of Proposition 2.

We start by restating the result that we need to prove:

**Proposition 2.** *Let  $\mathcal{P} \subseteq \mathbb{P}_n$  be a forest. Then there exists  $\eta \in (0, 1/2)$ , depending only on the degree  $d$ , such that for  $1 < p < \infty$  we have*

$$\|T^{\mathcal{P}}\|_p \lesssim_{p,d} 2^{-n\eta(1-\frac{1}{p^*})}.$$

#### 7.3.1. The $L^2$ bound.

Having in mind the procedures explained in Section 5.1. and Section 6.2., without loss of generality, we may assume that

$$\mathcal{P} = \bigcup_k \mathcal{P}_n^k,$$

such that each  $\mathcal{P}_n^k$  has the following properties:

- a) satisfies the construction described in Section 5.1. (we will preserve here all the notations from Section 5);
- b) is an  $L^\infty$ -**forest** (of  $n^{th}$  generation);
- c) any two trees inside the set are  $2^{100n}d$ -separated.

While properties a) and b) can be easily satisfied, for c) it is enough to discard from  $\mathcal{P}$  at most  $c(d)n$  families of negligible (incomparable) tiles for which we can apply Proposition 1.

Now let

$$\mathcal{A}_n^k := \left\{ P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \mid \begin{array}{l} I \subseteq A_n^k, A_{\mathbb{P} \setminus \bigcup_{j < n} \mathbb{P}_j, A_n^k}(P) \in [2^{-n}, 2^{-n+1}) \\ \text{if } J \in A_n^{k+1} \text{ s.t. } 20I \cap J^c \neq \emptyset \text{ then } |I| \geq |J| \end{array} \right\},$$

and

$$\mathcal{B}_n^k := \mathcal{P}_n^k \setminus \mathcal{A}_n^k.$$

Then, for each  $\mathcal{P}_n^k$ , we define a *boundary forest component*

$$\mathcal{P}_{n,bd}^k = \mathcal{P}_{n,bd}^{k,i} \cup \mathcal{P}_{n,bd}^{k,e}$$

where

$$\mathcal{P}_{n,bd}^{k,i} := \{P \in \mathcal{A}_n^k \mid \exists P_{kj} \in \mathcal{P}_n^{k,max} \text{ s.t. } P \leq P_{kj} \text{ and } 20I_P \cap (I_{P_{kj}})^c \neq \emptyset\}$$

and

$$\mathcal{P}_{n,bd}^{k,e} := \mathcal{P}_n^k \cap \mathcal{B}_n^k.$$

The *normal forest component* is defined as

$$\mathcal{P}_{n,nm}^k := \mathcal{P}_n^k \setminus \mathcal{P}_{n,bd}^k.$$

Finally, set  $\mathcal{P}_{bd} := \bigcup_k \mathcal{P}_{n,bd}^k$  and  $\mathcal{P}_{nm} := \bigcup_k \mathcal{P}_{n,nm}^k$ .

Now, here is our plan:

- for estimating the  $L^2$ -bound of the operator  $T^{\mathcal{P}_{nm}}$  we will show that the family  $\{T^{\mathcal{P}_{n,nm}^k}\}_k$  consists of almost orthogonal operators;
- for treating the operator  $T^{\mathcal{P}_{bd}}$  we just notice that  $\mathcal{P}_{bd}$  is a sparse forest and apply Proposition 1.

This being said, we claim that

$$(53) \quad \|T^{\mathcal{P}_{nm}}\|_2 \lesssim 2^{-\frac{n}{2}}.$$

For this it is enough to show that for  $|k - k'| > 10$  and some  $c > 1$  we have

$$(54) \quad \|T^{\mathcal{P}_{n,nm}^k} T^{\mathcal{P}_{n,nm}^{k'}}^*\|_2 \lesssim e^{-c|k-k'|n},$$

$$(55) \quad \|T^{\mathcal{P}_{n,nm}^k} T^{\mathcal{P}_{n,nm}^{k'}}\|_2 \lesssim e^{-c|k-k'|n}.$$

Indeed, (53) will then be easily derived, since

$$(56) \quad \|T^{\mathcal{P}_{n,nm}^{k'}} f\|_2 \lesssim 2^{-\frac{n}{2}} \|f\|_2.$$

Notice that (56) is a direct consequence of the Main Lemma since each  $\mathcal{P}_{n,nm}^k$  has the properties a), b) and c).

With this being said, let us start by proving (54).

Without loss of generality, we can suppose that  $k' \gg k$ . Applying Cauchy-Schwarz we have

$$\left| \left\langle T^{\mathcal{P}_{n,nm}^k} f, T^{\mathcal{P}_{n,nm}^{k'}} g \right\rangle \right| \leq \|\chi_{A_n^{k'}} T^{\mathcal{P}_{n,nm}^k} f\|_2 \|T^{\mathcal{P}_{n,nm}^{k'}} g\|_2.$$

Here we have used that  $\mathcal{P}_{n,nm}^{k'}$  is normal and thus  $\text{supp } T^{\mathcal{P}_{n,nm}^{k'}}^* \subseteq A_n^{k'}$ .

Next, from the way in which we have constructed  $\mathcal{P}_{n,nm}^k$ , we have that

$$\forall P \in \mathcal{P}_{n,nm}^k \text{ s.t. } I_{P^*} \cap A_n^{k+1} \neq \emptyset \Rightarrow I_{P^*} \not\subseteq A_n^{k+1}.$$

Thus, for any  $P \in \mathcal{P}_{n,nm}^k$ , we either have  $I_{P^*} \cap A_n^{k+1} = \emptyset$  or the following relation holds:

$$(57) \quad \frac{|I_{P^*} \cap A_n^{k'}|}{|I_{P^*}|} \leq \frac{|I_{P^*} \cap A_n^{k'}|}{|I_{P^*} \cap A_n^{k+1}|} \lesssim e^{-c|k-k'|n}.$$

Reached in this point, we remember that  $\mathcal{P}_{n,nm}^k$  is an  $L^\infty$ -forest of  $n^{\text{th}}$  generation and hence

$$(58) \quad \mathcal{P}_{n,nm}^k = \bigcup_{j=1}^{cn2^n} \mathcal{R}_j^k,$$

with each  $\mathcal{R}_j^k$  a row.

Then, using (57), and applying Lemma 4 for  $A := A_n^{k'}$ , we obtain

$$(59) \quad \left\| \chi_{A_n^{k'}} T^{\mathcal{P}_{n,nm}^k} f \right\|_2 \lesssim \sum_{j=1}^{cn2^n} \left\| \chi_{A_n^{k'}} T^{\mathcal{R}_j^k} f \right\|_2 \lesssim e^{-c|k-k'|n} \|f\|_2 ,$$

which proves (54).

Will pass now to the proof of (55).

As before, we can start by first applying Cauchy-Schwartz

$$\left| \left\langle T^{\mathcal{P}_{n,nm}^k} f, T^{\mathcal{P}_{n,nm}^{k'}} g \right\rangle \right| \leq \| \chi_{A_n^{k'}} T^{\mathcal{P}_{n,nm}^k} f \|_2 \| T^{\mathcal{P}_{n,nm}^{k'}} g \|_2 .$$

Based on (58) and the fact that the operators  $\{T^{\mathcal{R}_j^k}\}_j$  have disjoint supports, we have

$$\| \chi_{A_n^{k'}} T^{\mathcal{P}_{n,nm}^k} f \|_2^2 = \sum_j \| \chi_{A_n^{k'}} T^{\mathcal{R}_j^k} f \|_2^2 \lesssim n 2^n \sup_j \| \chi_{A_n^{k'}} T^{\mathcal{R}_j^k} f \|_2^2 .$$

Now applying Lemma 6 to our row  $\mathcal{R}_j^k$  (with the obvious adaptation of the partition  $\mathcal{J}_{\mathcal{P}}$  there to our new context - call this new partition  $\mathcal{J}_{\mathcal{R}_j^k}$ ) we have

$$(60) \quad \| \chi_{A_n^{k'}} T^{\mathcal{R}_j^k} f \|_2 \lesssim \left( \sup_{J \in \mathcal{J}_{\mathcal{R}_j^k}} \frac{|E_{A_n^{k'}}(J)|}{|J|} \right)^{\frac{1}{2}} \|f\|_2 .$$

Here, it is easy to remark that, from the construction of  $\mathcal{P}_{n,nm}^k$ , we have

$$\sup_{J \in \mathcal{J}_{\mathcal{R}_j^k}} \frac{|E_{A_n^{k'}}(J)|}{|J|} \lesssim e^{-c|k-k'|n} .$$

Thus, combining this last observation with (60), we deduce

$$\| \chi_{A_n^{k'}} T^{\mathcal{P}_{n,nm}^k} f \|_2 \lesssim n^{\frac{1}{2}} 2^{\frac{n}{2}} e^{-c|k-k'|n} \|f\|_2 ,$$

which together with (56) implies (55).

### 7.3.2. The $L^p$ bound.

In this section, based on the assumptions a), b) and c) made at the beginning of the proof, we will show that

$$(61) \quad \|T^{\mathcal{P}_{nm}}\|_p \lesssim_p 2^{-n(1-\frac{1}{p^*})} .$$

Heuristically, our first step will be to prove that for any  $1 < p < \infty$  we have

$$(62) \quad \left\| \sum_k T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p \lesssim_p \sum_k \|T^{\mathcal{P}_{n,nm}^k} f\|_p^p + \text{Error} ,$$

where the "Error" term above will be made precise in what follows.

Let us first take  $p$  to be a fixed (large) positive even integer.

Then, we notice that (up to conjugation), we have

$$\left\| \sum_k T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p \approx_p \sum_{\substack{(k_1, \dots, k_p), (r_1, \dots, r_p) \in \mathbb{N}^p \\ r_1 + \dots + r_p = p}} \int (T^{\mathcal{P}_{n,nm}^{k_1}} f)^{r_1} \dots (T^{\mathcal{P}_{n,nm}^{k_p}} f)^{r_p}$$

and after applying the Hölder and Jensen inequalities we further have

$$\begin{aligned} & \left\| \sum_k T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p \lesssim_p \\ & \sum_{\substack{(k_1, \dots, k_p), (r_1, \dots, r_p) \in \mathbb{N}^p \\ r_1 + \dots + r_p = p}} \left( \int_{\bigcap_{j=1}^p A_n^{k_j}} \left| T^{\mathcal{P}_{n,nm}^{k_1}} f \right|^p \right)^{\frac{r_1}{p}} \dots \left( \int_{\bigcap_{j=1}^p A_n^{k_j}} \left| T^{\mathcal{P}_{n,nm}^{k_p}} f \right|^p \right)^{\frac{r_p}{p}} \\ & \lesssim_p \sum_k \sum_{m \in \mathbb{N}} |m+1|^{100p} \int_{A_n^{k+m}} \left| T^{\mathcal{P}_{n,nm}^k} f \right|^p. \end{aligned}$$

Thus, we have just proved that for  $p \in 2\mathbb{N}$ , with  $p > 1$ , we have that

$$(63) \quad \left\| \sum_k T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p \lesssim_p \sum_k \sum_{m \in \mathbb{N}} |m+1|^{100p} \int_{A_n^{k+m}} \left| T^{\mathcal{P}_{n,nm}^k} f \right|^p.$$

On the other hand, by triangle inequality we trivially have

$$(64) \quad \left\| \sum_k T^{\mathcal{P}_{n,nm}^k} f \right\|_1 \lesssim \sum_k \left\| T^{\mathcal{P}_{n,nm}^k} f \right\|_1,$$

Using now interpolation between  $p = 1$  case and  $p \in 2\mathbb{N}$  ( $p > 1$ ) case, we conclude that for any  $1 < p < \infty$  the following holds

$$(65) \quad \left\| \sum_k T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p \lesssim_p \sum_k \left\| T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p + \sum_{\substack{m \geq 10p \\ m \in \mathbb{N}}} \sum_k m^{100p} \int_{A_n^{k+m}} \left| T^{\mathcal{P}_{n,nm}^k} f \right|^p.$$

Notice that (65) is the precise formulation of the heuristic described by (62).

Next step will be to treat the main term

$$A = \sum_k \left\| T^{\mathcal{P}_{n,nm}^k} f \right\|_p^p.$$

Firstly, applying Main Lemma, we have that

$$(66) \quad \left\| T^{\mathcal{P}_{n,nm}^k} f \right\|_2^2 \lesssim \sum_j \left\| T^{\mathcal{R}_j^k} f \right\|_2^2.$$

Interpolating this with the obvious triangle inequality in  $L^1$ , and respectively  $L^\infty$ , we deduce that for  $1 < p < \infty$  we have

$$(67) \quad \left\| T^{\mathcal{P}_{n,nm}^k} f \right\|_p \lesssim_p \left\{ \sum_j \left\| T^{\mathcal{R}_j^k} f \right\|_p^{p^*} \right\}^{\frac{1}{p^*}} \lesssim (n 2^n)^{\frac{1}{p^*} - \frac{1}{p}} \left\{ \sum_j \left\| T^{\mathcal{R}_j^k} f \right\|_p^p \right\}^{\frac{1}{p}}.$$



Next, for  $j, k$  fixed, we have

$$(68) \quad \|T^{\mathcal{R}_j^{k*}} f\|_p \lesssim \sum_{l \in \mathbb{N}} \|T^{\mathcal{R}_j^{k*}} \chi_{A_n^{k+l} \setminus A_n^{k+l+1}} f\|_p.$$

Further, from (57) and Lemma 6, we deduce

$$(69) \quad \|T^{\mathcal{R}_j^{k*}} (\chi_{A_n^{k+l} \setminus A_n^{k+l+1}} \cdot)\|_p = \|\chi_{A_n^{k+l} \setminus A_n^{k+l+1}} T^{\mathcal{R}_j^k}\|_{p'} \lesssim \min(2^{-\frac{ln}{p'}}, 2^{-\frac{n}{p'}}).$$

Denoting now with  $E_j^k := \bigcup_{P \in \mathcal{R}_j^k} E(P)$  and combining (68) and (69) we have

$$(70) \quad \|T^{\mathcal{R}_j^{k*}} f\|_p \lesssim \sum_{l \in \mathbb{N}} \min(2^{-\frac{ln}{p'}}, 2^{-\frac{n}{p'}}) \|\chi_{A_n^{k+l} \setminus A_n^{k+l+1}} \chi_{E_j^k} f\|_p,$$

which by another application of Hölder gives

$$(71) \quad \|T^{\mathcal{R}_j^{k*}} f\|_p \lesssim_p 2^{-\frac{n}{p'}} \left\{ \sum_{l \in \mathbb{N}} (l+1)^p \min(2^{-\frac{(l-1)n}{p'}}, 1) \|\chi_{A_n^{k+l} \setminus A_n^{k+l+1}} \chi_{E_j^k} f\|_p^p \right\}^{\frac{1}{p}}.$$

Combining now (67) and (71) we have

$$A \lesssim_p (2^n)^{p(\frac{1}{p^*} - \frac{1}{p})} 2^{-p \frac{n}{p'}} \times \sum_k \sum_j \sum_l (l+1)^p \min(2^{-\frac{(l-1)n}{p'}}, 1) \|\chi_{A_n^{k+l} \setminus A_n^{k+l+1}} \chi_{E_j^k} f\|_p^p.$$

If we add the fact that  $\{E_j^k\}_j$  are disjoint for each  $k \in \mathbb{N}$  we conclude

$$(72) \quad A \lesssim_p 2^{-n} \|f\|_p^p.$$

We pass now to the error term

$$B := \sum_{m \geq 10p} \sum_k m^{100p} \int_{A_n^{k+m}} |T^{\mathcal{P}_{n,nm}^{k*}} f|^p.$$

We first notice that

$$(73) \quad \int_{A_n^{k+m}} |T^{\mathcal{P}_{n,nm}^{k*}} f|^p \lesssim (n 2^n)^p \sum_j \int_{A_n^{k+m}} |T^{\mathcal{R}_j^{k*}} f|^p.$$

Now, based on (57) and the Lemma 7, we deduce that for each  $j$  we have

$$(74) \quad \int_{A_n^{k+m}} |T^{\mathcal{R}_j^{k*}} f|^p \lesssim 2^{-mn} \|f\|_p^p.$$

Combining (69) with (74) we further have

$$(75) \quad \int_{A_n^{k+m}} |T^{\mathcal{R}_j^{k*}} \chi_{A_n^{k+l} \setminus A_n^{k+l+1}} f|^p \lesssim_p 2^{-\frac{p(l+1)n}{4p'}} 2^{-\frac{mn}{2}} \|\chi_{A_n^{k+l} \setminus A_n^{k+l+1}} f\|_p^p.$$

Next, proceeding as in (70) and (71), we have

$$(76) \quad \int_{A_n^{k+m}} \left| T^{\mathcal{R}_j^{k*}} f \right|^p \lesssim_p 2^{-\frac{m \cdot n}{2}} \sum_{l \in \mathbb{N}} (l+1)^p 2^{-\frac{l \cdot n \cdot p}{4p}} \left\| \chi_{A_n^{k+l} \setminus A_n^{k+l+1}} \chi_{E_j^k} f \right\|_p^p.$$

Putting together (73) and (76) we deduce that

$$B \lesssim \sum_{m \geq 10p} m^{100p} 2^{-\frac{m \cdot n}{2}} (n 2^n)^p \sum_k \sum_{l \in \mathbb{N}} (l+1)^p 2^{-\frac{l \cdot n \cdot p}{4p}} \left\| \chi_{A_n^{k+l} \setminus A_n^{k+l+1}} f \right\|_p^p,$$

and hence

$$(77) \quad B \lesssim_p 2^{-n} \|f\|_p^p.$$

Finally, from (72) and (77), we conclude that (61) holds.  $\square$

## 8. Remarks

In this section we will discuss some applications and consequences of the discretization procedure presented in Section 5.1.

1) The first remark is a consequence of a fruitful conversation that I have had with C. Thiele and M. Bateman, and refers to a vector-valued variant of the Carleson Theorem. More precisely, using the above discretization procedure (and thus eliminating the exceptional sets), we devised an alternative proof that for any  $1 < p, q < \infty$  one has <sup>16</sup>

$$(78) \quad \left\| \left( \sum_k |C_1 f_k|^q \right)^{\frac{1}{q}} \right\|_p \lesssim_{p,q} \left\| \left( \sum_k |f_k|^q \right)^{\frac{1}{q}} \right\|_p,$$

an inequality that had been proven in [5] using weighted and extrapolation theory. Nevertheless, as a consequence of the Theorem presented in this paper, one has that (78) holds with  $C_1$  replaced by  $C_d$ .

2) As mentioned in the introduction, our discretization procedure was designed for obtaining the following informal principle:

If  $\mathcal{P} = \bigcup_k \mathcal{P}_k \subseteq \mathbb{P}_n$  is a collection of separated trees, then

$$(79) \quad \left\| \sum_k T^{\mathcal{P}_k^*} f \right\|_2 \lesssim \log(10 + \|\mathcal{N}_{\mathcal{P}}\|_{BMO_C}) \left( \sum_k \|T^{\mathcal{P}_k^*} f\|_2^2 \right)^{\frac{1}{2}}.$$

In both [3] and [4], this principle was only present in the weaker form with  $\|\mathcal{N}_{\mathcal{P}}\|_{BMO_C}$  replaced by  $\|\mathcal{N}_{\mathcal{P}}\|_{L^\infty}$ , thus causing some further expense of work in carefully treating the sets on which  $\|\mathcal{N}_{\mathcal{P}}\|_{L^\infty}$  was too large. Through (79) we can now avoid all these technicalities.

3) Finally, the last remark deals with the behavior of the Carleson operator  $C$  near  $L^1$ . The question in discussion here is:

---

<sup>16</sup>Here we use the notations from Section 1.

What is the maximal rearrangement-invariant space  $Y \subset L^1$  for which the following relation

$$(80) \quad \|Cf\|_{1,\infty} \lesssim \|f\|_Y$$

holds for all  $f \in Y$ ?

In [10], using Carleson's original construction from [2] together with an  $L^p$  estimate for the level sets of the operator  $C$  inspired by the work of Hunt in [6], Sjölin proved that one may take in (80) the space  $Y = L \log L \log \log L$ . Moreover, he also showed that

$$(81) \quad \|Cf\|_1 \lesssim \|f\|_{L(\log L)^2}.$$

On the other hand, the original approach of Fefferman was confined to showing that (80) holds for  $Y = L(\log L)^M$  for some large  $M > 2$ .

Later on, using Banach rearrangement techniques and relying on Sjölin's proof, Y. Antonov showed in [1] that one can enlarge the  $Y$ -space to  $L \log L \log \log L$  which remains the best current result to date.

For the time being, in the absence of the exceptional sets, we can use our construction to reprove (81) by different means. Moreover, we are able to show that by a slight modification of the definition of the tile families  $\{\mathbb{P}_n\}$  we have that for each  $n \in \mathbb{N}$  the following holds:

$$(82) \quad \left\| T^{\mathbb{P}_n} f \right\|_1 \lesssim \|f\|_{L \log L}.$$

While it is not possible to proceed as in the  $L^p$  ( $1 < p < \infty$ ) case by summing up the components  $\|T^{\mathbb{P}_n} f\|_1$  (since there is no decay in the mass parameter  $n$ ), we hope that one may be able to apply a further regrouping of the tiles, possibly depending on a second parameter which takes into account the structure (size and localization) of the function  $f$ , and show an improved bound for the quantity  $\left\| \sum_n T^{\mathbb{P}_n} f \right\|_1$ .

As one can see, the natural question in discussion is how close can one push the space  $Y$  in (80) to the "ideal" limiting space  $L \log L$ .

This interesting problem remains open for further investigations.

## 9. Appendix - Results on the $L^\infty$ -distribution of polynomials

**Lemma A.** *If  $q \in \mathcal{Q}_{d-1}$  and  $I, J$  are some (not necessarily dyadic) intervals obeying  $I \supseteq J$ , then there exists a constant  $c(d)$  such that*

$$\|q\|_{L^\infty(I)} \leq c(d) \left( \frac{|I|}{|J|} \right)^{d-1} \|q\|_{L^\infty(J)}.$$

*Proof.* Let  $\{x_J^k\}_{k \in \{1, \dots, d\}}$  be obtained as in the procedure described in Section

2. Then, since  $q \in \mathcal{Q}_{d-1}$ , for any  $x \in I$  we have that

$$q(x) := \sum_{j=1}^d \frac{\prod_{\substack{k=1 \\ k \neq j}}^d (x - x_J^k)}{\prod_{\substack{k=1 \\ k \neq j}}^d (x_J^j - x_J^k)} q(x_J^j).$$

As a consequence,

$$\|q\|_{L^\infty(I)} \leq d \|q\|_{L^\infty(J)} \sup_{\substack{j \\ x \in I}} \left| \frac{\prod_{\substack{k=1 \\ k \neq j}}^d (x - x_J^k)}{\prod_{\substack{k=1 \\ k \neq j}}^d (x_J^j - x_J^k)} \right| \leq d \|q\|_{L^\infty(J)} \frac{|I|^{d-1}}{(|J|/d)^{d-1}}.$$

□

**Lemma B.** *If  $q \in \mathcal{Q}_{d-1}$ ,  $\eta > 0$  and  $I \subset \mathbb{T}$  some (dyadic) interval, then*

$$(83) \quad |\{y \in I \mid |q(y)| < \eta\}| \leq c(d) \left( \frac{\eta}{\|q\|_{L^\infty(I)}} \right)^{\frac{1}{d-1}} |I|.$$

*Proof.* The set  $A_\eta = \{y \in I \mid |q(y)| < \eta\}$  is the pre-image of  $(-\eta, \eta)$  under a polynomial of degree  $d-1$ , so it can be written as

$$A_\eta = \bigcup_{k=1}^r J_k(\eta),$$

where  $r \in \mathbb{N}$ ,  $r \leq d-1$  and  $\{J_k(\eta)\}_k$  are open intervals. Now all that remains is to apply the previous lemma with  $J = J_k(\eta)$  for each  $k$ .

□

**Lemma C.** *If  $P = [\alpha^1, \alpha^2, \dots, \alpha^d, I] \in \mathbb{P}$  and  $q \in P$ , then*

$$\|q - q_P\|_{L^\infty(\tilde{I})} \leq c(d) |I|^{-1}.$$

*Proof.* Set  $u := q - q_P$ ; then, since both  $q, q_P \in P$ , we deduce (for all  $k \in \{1, \dots, d\}$ ):

$$u(x_I^k) \in [-|I|^{-1}, |I|^{-1}].$$

On the other hand,

$$u(x) := \sum_{j=1}^d \frac{\prod_{\substack{k=1 \\ k \neq j}}^d (x - x_I^k)}{\prod_{\substack{k=1 \\ k \neq j}}^d (x_I^j - x_I^k)} u(x_I^j) \quad \forall x \in I.$$

Then, proceeding as in Lemma A, we conclude

$$\|u\|_{L^\infty(I)} \leq d |I|^{-1} \frac{|I|^{d-1}}{(|I|/d)^{d-1}} \leq d^d |I|^{-1}.$$

□

## REFERENCES

- [1] N. Yu. Antonov, *Convergence of Fourier Series*, East J. Approx. **2** (1996), no. 2, 187-196.
- [2] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135-157.
- [3] C. Demeter, T. Tao, C. Thiele, *Maximal multilinear operators*, Trans. Amer. Math. Soc. **360** (2008), no. 9, 4989-5042.
- [4] C. Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. **98** (1973), 551-571.
- [5] L. Grafakos, J. M. Martell, F. Soria, *Weighted norm inequalities for maximally modulated singular integral operators*, Math. Ann. **331** (2005), 359-394.
- [6] R. Hunt, *On the convergence of Fourier series*, Proceedings of the Conference on Orthogonal Expansions and their Continuous Analogues, Carbondale (1968), 235-255.
- [7] M. Lacey, *Carleson's theorem with quadratic phase functions*, Studia Math. **153** (2002), 249-267.
- [8] M. Lacey and C. Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361-370.
- [9] V. Lie, *The (weak- $L^2$ ) Boundedness of The Quadratic Carleson Operator*, Geom. Funct. Anal. **19** (2009), 457-497.
- [10] P. Sjölin, *An inequality of Paley and convergence a.e. of Walsh-Fourier series*, Ark. Mat. **7** (1969), 551-570.
- [11] E. M. Stein, *On limits of Sequences of Operators*, Ann. of Math. **74** (1961), 140-170.
- [12] E. M. Stein, *Oscillatory integrals related to Radon-like transforms*, Proceed. of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993), 1995, 535-551.
- [13] E. M. Stein and S. Wainger, *Oscillatory integrals related to Carleson's theorem*, Math. Res. Lett. **8** (2001), 789-800.

DEPARTMENT OF MATHEMATICS, PRINCETON, NJ 08544-1000 USA

*E-mail address:* `vlie@math.princeton.edu`

INSTITUTE OF MATHEMATICS OF THE ROMANIAN ACADEMY, BUCHAREST, RO 70700

P.O. BOX 1-764